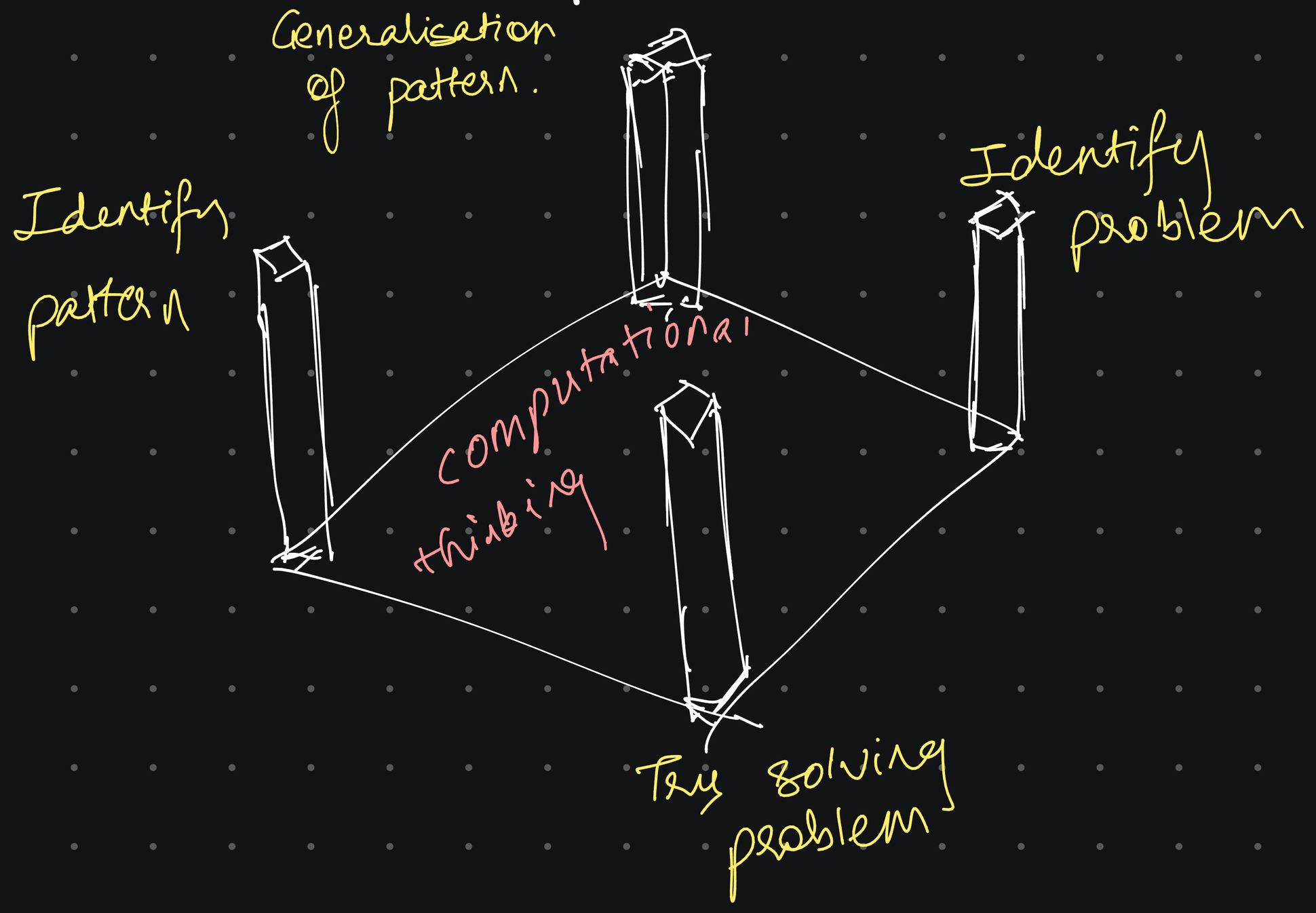
11/8

→ 4 pillars of computational thinking

1) Identify problem
2) Try solving problem
3) Identify pattern
4) Generalisation of pattern



14/8

Books

→ N G Das : Statistical Methods I & II

→ Gerald Keller

→ Gupta Kapoor Mathematical Statistics

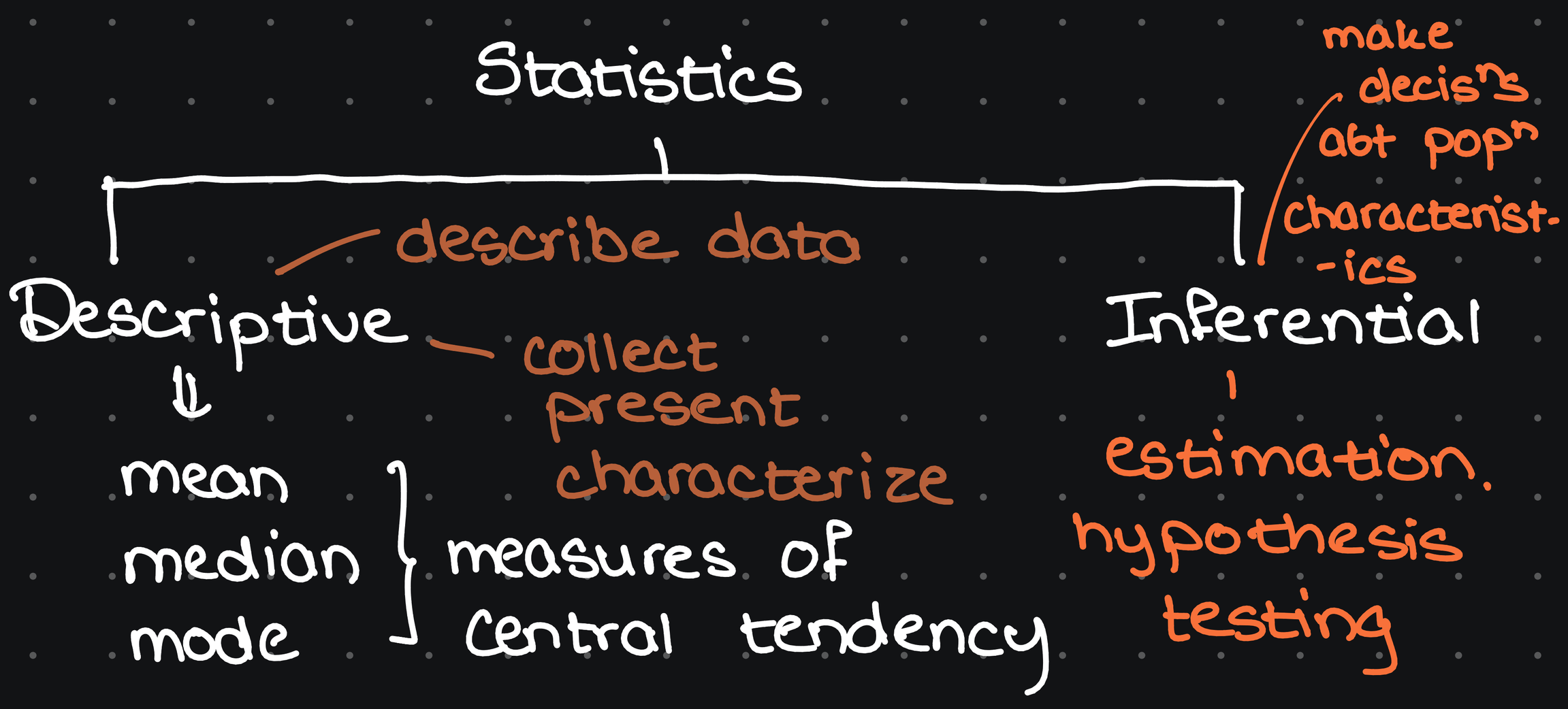
→ Statistical Methods I & II - Nanda Rajput

→ Anderson & Sweeney Williams

→ Mathematical Statistics w/ applications — Kapadia

→ Modern mathematical statistics with applicat^s pg 249-252 — Devore

Fundamentals of Statistics



" making sense of data "

→ What is stats? Collecting, Presenting and Characterizing data

1. Identify target population
2. Collecting data eg. survey
3. Presenting data eg. charts & tables
4. Characterizing data eg. average

Stats: science of data. Involves collecting, classifying, summarizing, organising, analyzing & interpreting numerical info.

Processes

- Describing sets of data
- Drawing conclusions about sets of data based on sampling
  - ↓
  - rely on estimation

Expected ⟷ Observed

margin of error = E - O

try to reduce error

Fundamental elements

1. Expt. unit
   object upon which we collect data

2. Populatn ——— Parameter
   * all items of interest

3. Variable
   characteristic of individual expt. unit

4. Sample ——— Statistic
   * subset of popn units

— Statistical Inference
estimate / predictn / generalizatn abt popn based on info collected in a sample

— Measure of Reliability
statement (usually qualified) abt degree of uncertainty associated with statistical inference

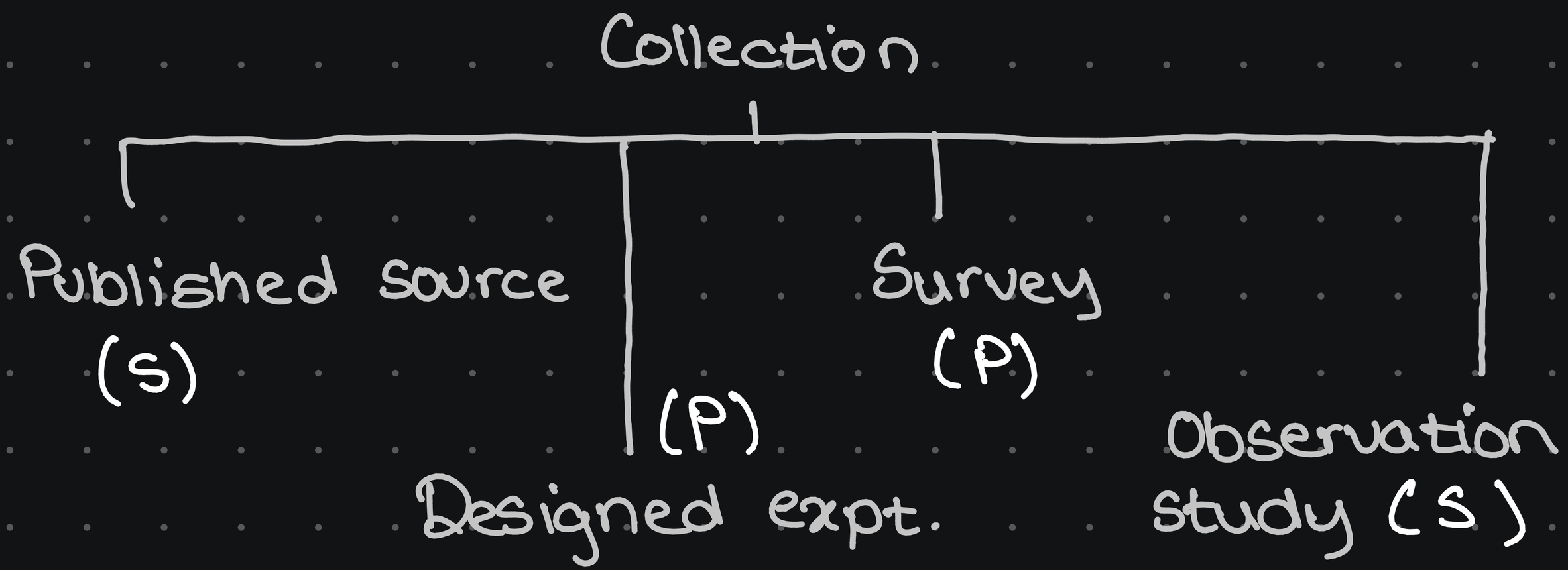
Elements of Descriptive Stats

1. Popn / sample of interest
2. Variables (characteristics)
3. Tables / graphs / numerical summary tools
4. Identificatn of patterns

Elements of Inferential Stats

1. Popn
2. Variables (characteristics)
3. Sample
4. Inference (abt popn based on sample)
5. Reliability (for inference)

Aru ❤ Rashe



Strict control
over units

Random Sample : every sample of size 'n'
has equal chance of selectⁿ

Statistical Thinking – applying rational thots
& science of stats to critically assess
data & inferences.

Fundamental : variatⁿ exists in popⁿ and
process data

Rationality is less & we have subconscious
bias due to various agendas (political,
economical, social etc.)

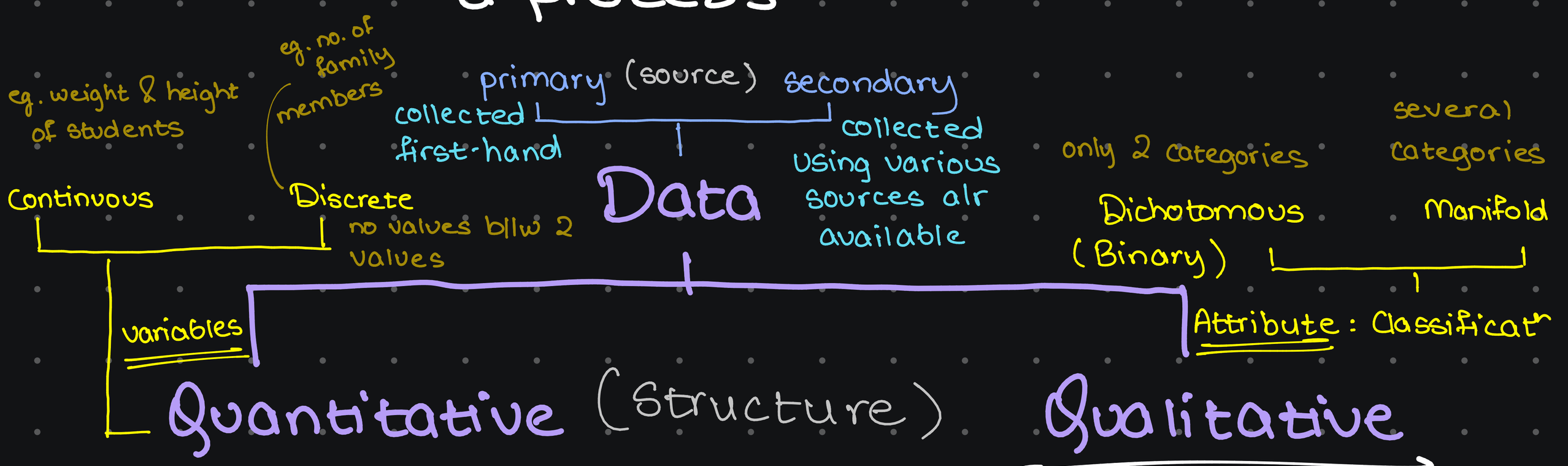
We should not get influenced by ANY
personal agendas

Process: series of acts/operats that transforms inputs to outputs

Produces/generates output over time

Black box → process whose operats/acts are unknown or unspecified

Sample → any set of output produced by a process



recorded on naturally occurring numerical scale

can be classified into one grp. of categories
cant be measured numerically

Rankings: Ordinal scale — scaling of data

Feasibility Analysis : whether our new venture will be successful or not?

→ geographical locatⁿ
↓
rates
ROI

choices of food

→ Target audience / Demography
→ Competitors
→ Pre existing branding - Franchise
→ Demand & supply
→ Menu
→ Footfall
→ Promotⁿ investment decisⁿ
→ ~~Facilities available~~
Infra

→ Delivery optⁿs

partners

+preferences

+transport

dine-in takeaway

→ cost of sourcing ingredients
→ employees & staff availability
→ habitatⁿs & habits
→ accessibility

Managerial formulatⁿ of problem

↓

Managerial questⁿ relating to problem

↓

Statistical formulatⁿ of questⁿ (hypothesis)

↘

Reformulatⁿ of problem

New q. ←

Answer to managerial questⁿ

↑

Managerial soln to problem

↑

Answer to statistical questⁿ

↗

STATISTICAL ANALYSIS

Econometrics
study eco & stats 2gether

take decisⁿ and do forecasting

↳ create mathematical model

21/8

Signal / Noise ⇒ effective data ratio

less the noise, better the signal

NOISE = Primary data + false info
non-respondent
inaccurate data
outliers/errors
bias
judgemental data

→ role of filtering

if filtering becomes too time consuming, go back to data collection method ↓

SIGNAL

sampling: reduce noise in the data to get max. possible accurate data

Popn / Universe: aggregate from which sample is taken

- homogenous: same everywhere
  eg. blood test
  blood of hand & leg is same
- heterogenous: different

SAMPLING

Sampling frame : list from which potential respondents are drawn

Sample : smaller collect^{n} of units from a popn used to define truths abt. that popn

Theoretical popn — who do you want to generalize to?

↓

Study popn — what popn can you get access to?

↓

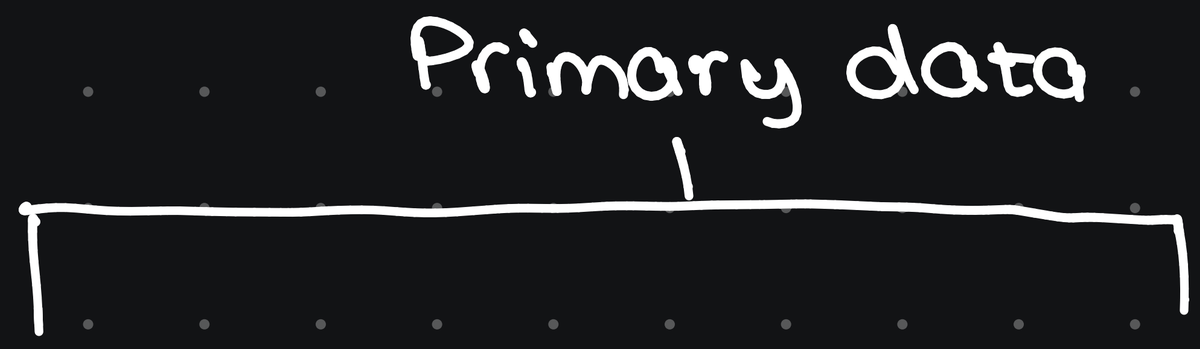
Sampling frame — how can you access?

↓

Sample — who is in your study?

Sampling : process of learning abt popn on basis of sample drawn

elements: selecting sample
collecting info
making inference



→ data from each &
every unit
→ basis of various
surveys
→ results more
representative.
reliable, accurate

→ more effort, time
& money
→ big problem in
underdup. countries

advantages of sampling:

1. less resources
2. less workload
3. Give results with known accuracy
that can be calculated.

1. Law of Statistical Regularity

Sample is taken at random from a popⁿ, it is likely to possess same characteri-stic as that of popⁿ

2. Law of inertia of large numbers

Larger the size of sample, more accurate the results are likely to be

Sampling Process

1. Define target popⁿ
↓
2. Specify sampling frame (items/ events possible to measure)
↓
3. Specify sampling method
↓
4. determine sample size
↓
5. implement sampling plan
↓
6. data collectⁿ

Essentials of Sampling:

1. Representativeness : random selectⁿ
2. Adequacy : sample size
3. Independence : same selectⁿ chance
4. Homogeneity : no basic difference in nature of units

Sampling Methods

non-random	random	
Non-probability	**Probability**	**Mixed**
judgement	simple random	multistage
	(lottery method, random no. tables)	
quota	stratified random	multiphase
conference	systematic	lot quality assurance
snowball	cluster	

1] JUDGEMENT SAMPLING

→ purposive / deliberate

→ depends exclusively on judgement of investigator

⇓

selects most typical of the universe sampling.

→ small no. of sampling units
→ study unknown traits / case sampling
→ urgent public policy & business decisions

→ personal prejudice & bias
→ no objective way of evaluating reliability of results

2] Convenience Sampling

→ convenient sample units selected neither by probability nor judgement

→ useful in pilot studies

→ results biased & unsatisfactory

3] Quota Sampling

→ quotas set up according to some specified characteristic

→ within quota, selectⁿ depends on personal judgement

→ personal prejudice & bias

4] SNOWBALL SAMPLING

→ used when desired sample characteristic is rare

⇓

diff. & cost prohibitive to locate respondents

→ relies on referrals from initial subjects to generate additnal subjects

→ Steps:

Make contact with 1/2 cases in popn
Ask cases to identify further cases
Ask new cases to identify further new cases

→ access to diff. to reach popn

→ not representative of sample & will result in biased sample (self-selecting)

i] Simple Random

→ each unit has equal opportunity of being selected
chance determines which items shall be included

→ characteristics:

1. all items selected independently
2. at each selectⁿ, all remaining items have same chance of being selected
3. all possible samples of a given size are equally likely to be selected

→ no personal bias
sample more representative of popⁿ
accuracy can be assessed as sampling errors follow principles of chance

→ requires completely catalogued universe
cases too widely dispersed - more time & cost

2] Stratified random

→ issues:

base of stratificatn
no. of strata
sample size within strata

- proportal: propornⁿ in each stratum
- disproportal: equal no. in each stratum

→ more representative
greater accuracy
greater geographical concentratn

→ utmost care in dividing strata
skilled sampling supervisors
cost per observatn is high

3] Systematic Sampling

→ select first unit at random
select additional units at evenly spaced intervals

$k = \frac{N}{n}$ → Universe size
n → Sample size

↓
sampling interval

example

→ simple, convenient
less time consuming

→ pop^n with hidden periodicities ← hidden pattern

Stratification		Clustering
all strata represented in sample		only subset of clusters in sample
less error compared to simple random		more error compared to simple random
more expensive to obtain stratificatn info before sampling		reduces costs to sample only some areas / organisatns

4] Cluster

→ entire popn of interest divided into grps / clusters & random sample of clusters selected)

mutually exclusive + include entire popn

→ no units from non-selected clusters included

→ Clusters ⇓ Units — primary sampling units / secondary sampling units

Sampling interval = $\frac{\text{cumulative popn}}{\text{no. of popn units}}$

1st cluster + S.i. = 2nd cluster

eg.

	Freq	c.f.	Cluster	si = 24000/8
I	2000	2000	1	
II	3000	5000	2	= 3000
III	2500	7500		
IV	4000	11500	3	
V	5000	16500	4, 5	
VI	2500	19000		
VII	2000	21000	6	
VIII	3000	24000	7	

→ most economical
larger sample for similar fixed cost
less time for listing & implementatn
reduce travel & administrative costs

→ may not reflect community diversity
standard errors high

1] Multistage

→ carried out in various stages

→ banks on multiple randomizat°s

→ used when complete list of all members of popn doesn't exist / is inappropriate

→ flexibility
cover large area
valuable in under-dvpd. areas
enables existing divisions & sub-divisions of popn to be used as units

→ less accurate than sample chosen by single stage process

2] Multiphase

→ studies to be carried out in multiple phases

3] Lot quality assurance

for quality control

outcomes: acceptable
not acceptable

→ decision value: no. of defective items that need to be found before lot is deemed unacceptable

Risks

Type - I error
risk of accepting 'bad' lot

Type - II error
risk of not accepting 'good' lot

→ popⁿ divided into sets of non-overlapping lots
  └ serve same purpose as strata
  └ samples taken from every lot & proportⁿ of defectives in each lot calc.

binary response (acceptable/not)
  └ can use small sample size
} adv. over traditⁿal

Lottery method

With replacement

$P(x) = \frac{1}{N}$

every trial is a fresh trial

constant probability

Without replacement

$P(1) = \frac{1}{N}$

$P(2) = \frac{1}{N-1}$

↓

"trunketed sample space"

reduced

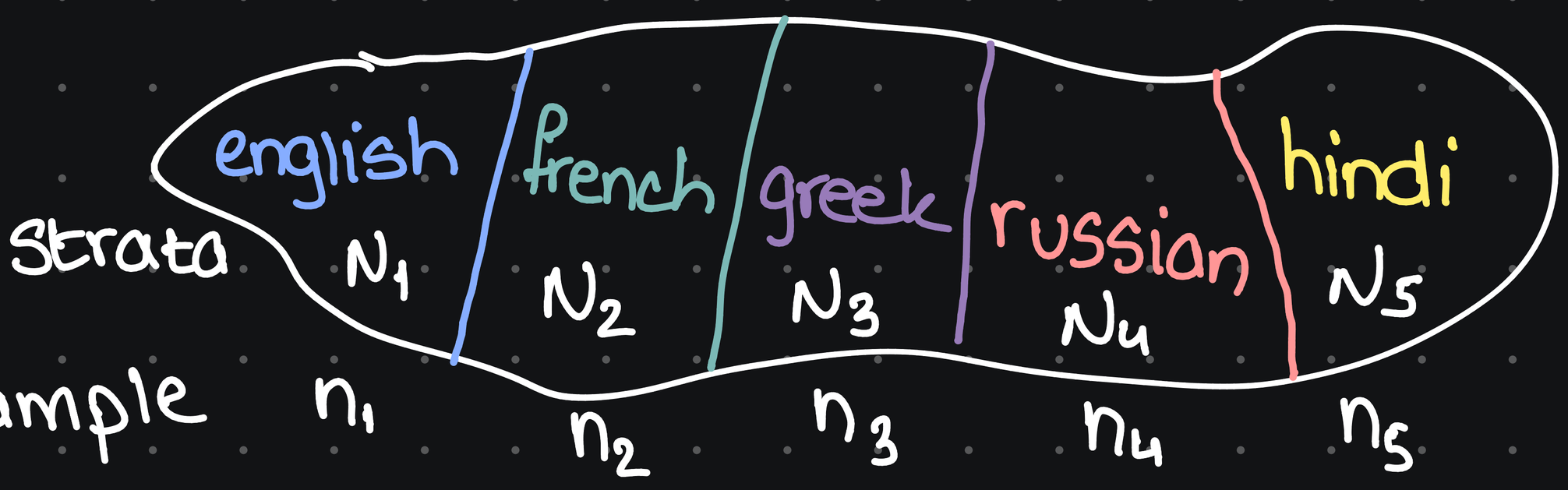
"conditional probability"

* Stratified Random Sampling

universe divided into mutually exclusive groups

simple random sample chosen indepen-dently from each group

sub sample size $\propto$ strata $n_i \propto N_i$



NOTES

1. Gerald Keller

Statistics for management | ch 5 | pg 186

Methods of collecting data

response rate: proportⁿ of all people who were selected who complete the survey
low: destroys validity

- Direct Observatⁿ
  - "observational data"
  - simplest method
  - difficult to produce useful info.
  - relatively inexpensive
- Experi-ments
  - "experimental"
  - expensive but better
- Surveys
  - └ personal interview
    - interviewer soliciting info from a respondent by asking prepared questions
  - └ telephone interview
  - └ self-administered survey
    - mailed to people.

└ questionnaire design

1. as short as possible
2. as short, simple & clearly worded (no ambiguity)
3. begin with simple demographic qs
4. dichotomous questions
5. avoid leading questions

Sampling

chief motive for examining sample rather than popⁿ is "cost."

sample propⁿ used as an **estimate**

target popⁿ : popⁿ abt. which we want to draw inferences

sampled popⁿ : actual popⁿ from which sample has been taken

self-selected samples almost always biased
↳ popⁿ composed entirely of people who are interested in the issue
eg. radio

SLOP (self-selected opinion poll)
Oy vey (Yiddish lament)

Sampling plans

Simple Random

every possible sample with the same no. of observatns is equally likely to be chosen

- SRSWR (with rep)
- SRSWOR (without rep)

Stratified Random

separating popn into mutually exclusive strata. & drawing simple random samples from each stream

Cluster

simple random sample of groups / clusters of elements

Errors

more serious because even larger sample size wont diminish size/possibility of occurrence

Sampling error
sample

differences b/tw sample & popn that exists only bcoz of the observatns that happened to be selected for the sample

- Biased
  - eliminatn of all sources of bias
- Unbiased
  - increase sample size

Non-Sampling error
sample & census

result from mistakes made in acquisition of data/ from sample observatns

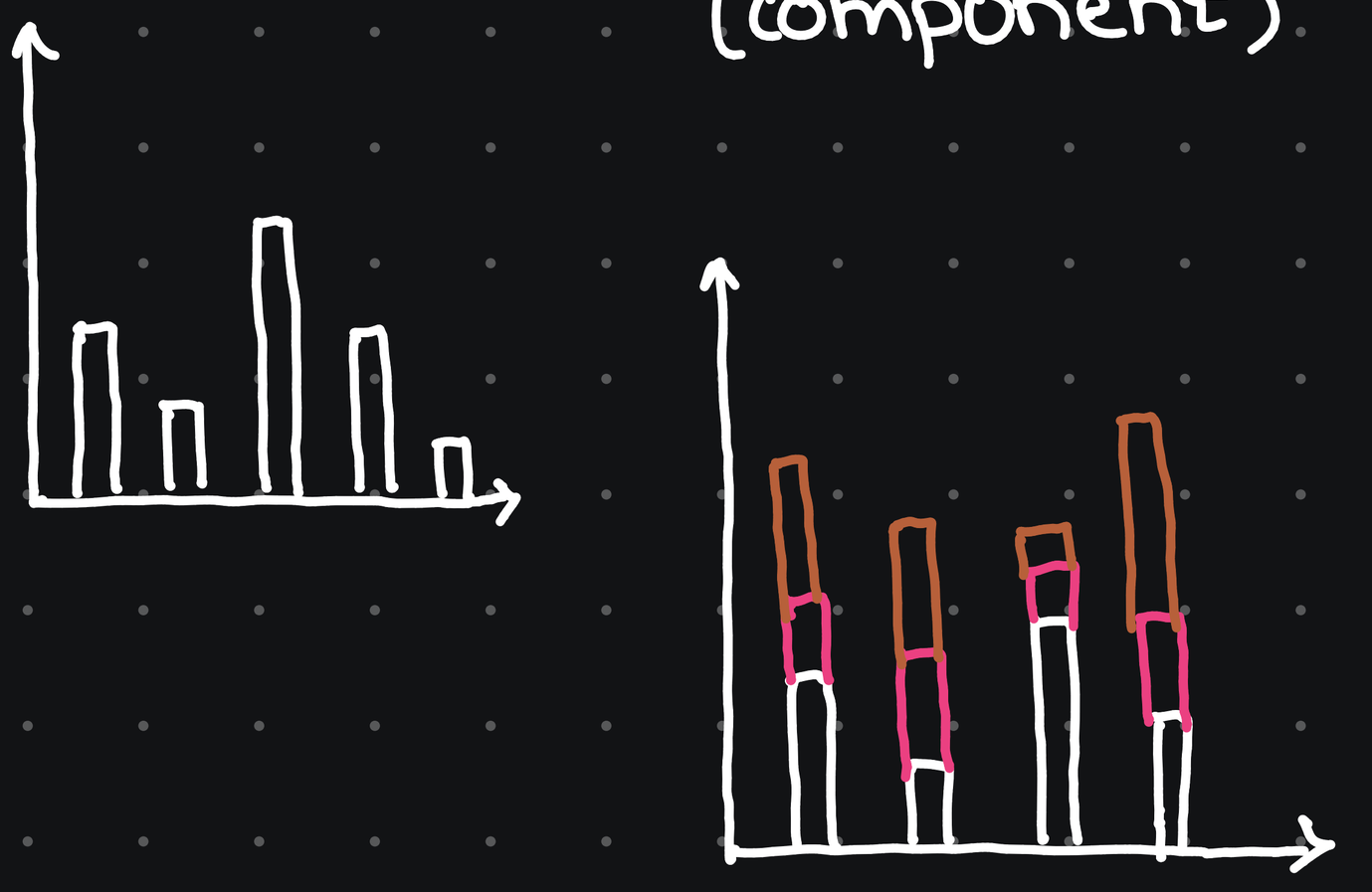
- Errors in acquisitn of data
  - recording incorrect responses
    - incorrect measurements (faulty equipment)
    - transcriptn mistakes
    - misinterpreted terms
    - sensitive issues
- Non-response error
  - responses not obtained from some members of sample
- Selectn bias
  - some members cannot possibly be selected for inclusn in sample

23/8

Diagrams...

① Bar Graphs

- Single
- Sub-divided (component)
- Multiple
- Percentage



② Histogram: for grouped data

no gap b/lw consecutive bars

area of bars proportional to fq.

* Frequency density = $\frac{\text{Frequency}}{\text{Class width}}$ } when class intervals are not equal

eg.

CI	fq.	fd.	rfq.	efq	error
20 - 30	34	3·4	0·15	45·2	
30 - 40	45	4·5	0·19	45·2	
40 - 50	72	7·2	0·31	45·2	
50 - 70	54	2·7	0·23	45·2	
70 - 100	21	0·7	0·09	45·2	
	226		1		

*

Relative frequency = $\frac{\text{frequency}}{\text{total observaths}}$

expected frequency = $\frac{226}{5}$ = 45·2

assuming uniform distribution of the data

Dispersion : scatterness/variation of observat's from their average

Characteristics :

1. rigidly defined
2. based on all items
3. not unduly affected by extreme items
4. lend itself for algebraic manipulation
5. simple to understand
6. easy to calculate

1. Range

→ difference between largest & smallest values of variable

- **Merits**
  1. simple to understand
  2. easy to calculate
  3. used in problems of quality control, weather forecasts etc.

- **Demerits**
  1. affected by extremes
  2. based on only 2 extreme observat's
  3. can't be calculated from open-end C.I.
  4. not suitable for mathematical treatment
  5. rarely used

2. Standard Deviation

→ positive square-root of arithmetic mean of square of deviations of given observations from their arithmetic mean

- **Merits**

1. rigidly defined, definite value
2. based on arithmetic mean — all merits apply
3. most imp. & widely used
4. possible for further algebraic treatment
5. less affected by fluctuat^n of sampling
   ↳ stable
6. basis for measuring coeff. of correlat^n

★ SD does not change if the same number is subtracted from all observations ★

- **Demerits**

1. not easy to understand, diff. to calculate
2. gives more weight to extreme values (because squared up)
3. absolute measure of variability
   ↳ cannot be used for comparison

3. Variance (& coeff. of variation)

↓

square of
Standard deviatn

Coefficient of Variation

Greater

- more variable
- less stable
- less uniform
- less consistent
- less homogeneous

less

- less variable
- more stable
- more uniform
- more consistent
- more homogenous

28/8

Ungrouped data : individual observat's

eg. 121, 111, 123, 127, 118, 125, 133, 182, 130, 140, 143, 147, 152

	$\overline{x}$		$\overline{x} - a$
⇒	111		~~-19~~
	118		~~-12~~
	121		~~-9~~
	123		~~-7~~
	125		~~-5~~
	127		~~-3~~
	130	⟶ median	0
	133		~~3~~
	140		10
	143		~~13~~
	147		~~17~~
	152		~~22~~
	182		52
	1752		62

$\frac{1752}{13} = 134.76$ ⟶ mean

OR $a = 130$

$m = 130 + \frac{62}{13} = 130 + 4.76 = 134.76$

CI	x_i	f_i	$f_i x_i$	(<) cf	(>) cf
10 - 30	20	5	100	5	45
				+7	-5
30 - 50	40	7	280	12	40
median class				+11	-7
50 - 70	60	11	660	23	33
modal class					
70 - 90	80	13	1040	36	22
90 - 110	100	6	600	42	9
110 - 130	120	3	360	45	3
		45	3040		

$$\text{mean} = \frac{\Sigma f_i x_i}{\Sigma f_i} = \frac{3040}{45} = \frac{608}{9} = \boxed{67.55}$$

$$\frac{45}{2} = 22.5$$

$$\text{median} = l + \left[\frac{\frac{n}{2} - cf}{f}\right] \times h \qquad = 50 + \frac{210}{11}$$

l — of class above

f — of median class

$$= 50 + 19.09$$

$$= 50 + \left[\frac{22.5 - 12}{11}\right] \times 20$$

$$= \boxed{69.09}$$

$$= 50 + \frac{10.5}{11} \times 20$$

$$\text{mode} = l + \left(\frac{f_1 - f_0}{2f_1 - f_2 - f_0}\right) h$$

$$= 70 + \left(\frac{13 - 11}{26 - 11 - 6}\right) 20$$

$$= 70 + \left(\frac{2}{9}\right) 20$$

$$= 70 + \frac{40}{9}$$

$$= 70 + 4.44$$

$$= \boxed{74.44}$$

Cumulative frequency

less than (<)
(upper limit)

more than (>)
(lower limit)

$h = 50$ $\quad\quad$ $x_i/25$

CI	x_i	f_i	cf	u_i	$u_i f_i$	$f_i x_i$
100-150	125	8	8	5	$^{2}40$	1000
150-200	175	11	19	7	77	1925
200-250	225	13	32	9	$^{3}117$	2925
250-300	275	7	39	11	77	1925
300-350	325	5	44	13	65	1625
350-400	375	2	46	15	30	750
		46			406	10150

$$\text{median} = l + \left[\frac{n/2 - cf}{f}\right] h$$

$$= 200 + \left[\frac{23 - 19}{13}\right] 50$$

$$= 200 + \left[\frac{4}{13}\right] 50$$

$$= 200 + \frac{200}{13}$$

$$= 200 + 15.38$$

$$= \boxed{215.38}$$

$$\text{mode} = l + \left(\frac{f_1 - f_0}{2f_1 - f_0 - f_2}\right) h$$

$$= 200 + \left(\frac{13 - 11}{26 - 11 - 7}\right) 50$$

$$= 200 + \frac{2}{8} \times 50$$

$$= 200 + \frac{1}{4} \cdot 50$$

$$= 200 + 12.5$$

$$= \boxed{212.5}$$

$$\text{mean} = \frac{\Sigma u_i f_i}{\Sigma f_i} \times 25$$

$$= \frac{406}{46} \times 25$$

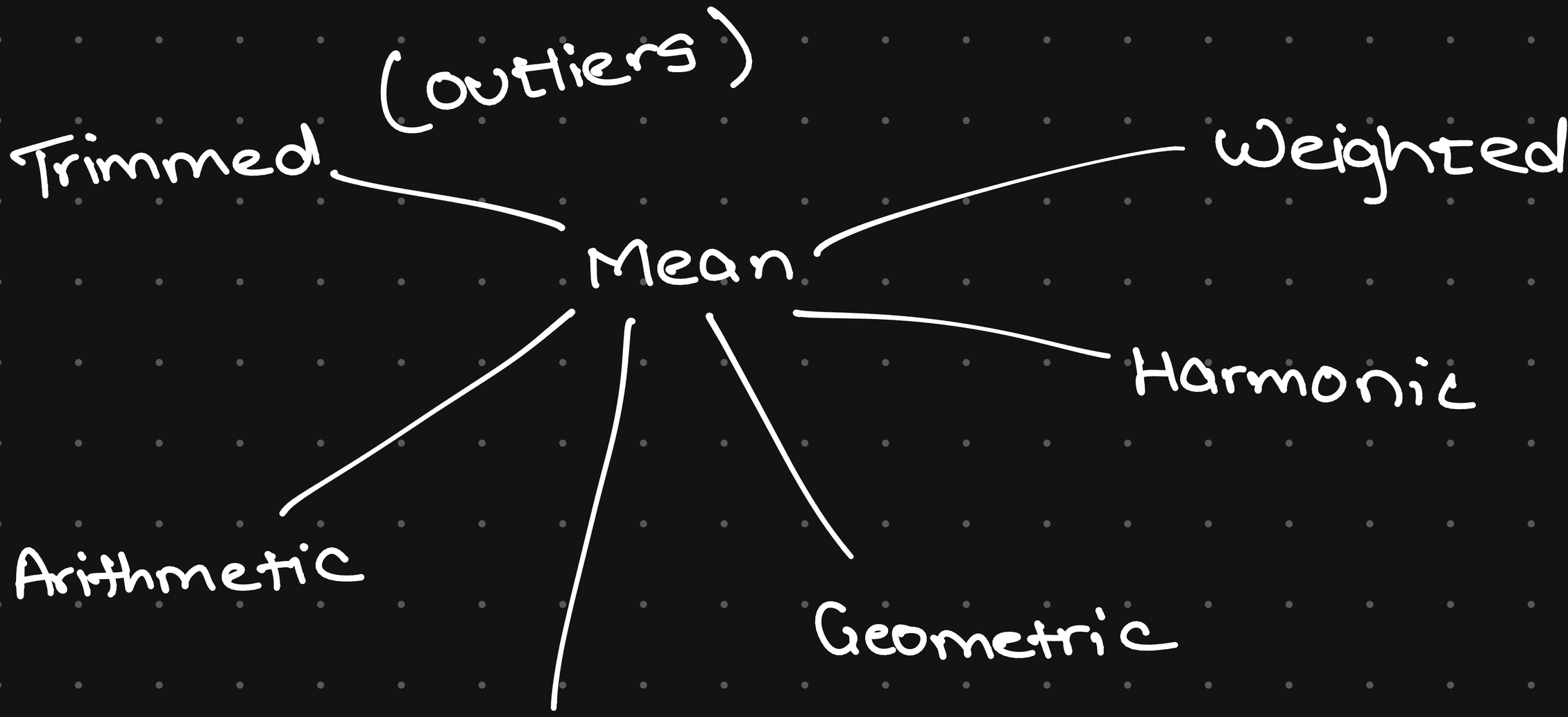
$$= 8.82 \times 25$$

$$= \boxed{220.5}$$

Empirical Relation

└ approximate

Mode ≈ 3 Median - 2 Mode



Grouped Mean

$$\bar{x} = \frac{\sum_{i=1}^{k} n_i \bar{x}_i}{\sum n_i}$$

mean annual salary = ₹25.000

Male : ₹27.000
Female : ₹17.000
} find percentage of male & female employees

$$\frac{27.000m + 17.000f}{m+f} = 25.000$$

$$27.000m + 17.000f = 25.000m + 25.000f$$

$$2.000m = 8.000f$$

$$\frac{m}{f} = \frac{8000}{2000} = \frac{4}{1}$$

80% m
20% f

$$\frac{m}{f+m} = \frac{4}{5} = 80\%$$

• : grouped
• : ungrouped

Measures of Central Tendency

Mean	Median	Mode
$\frac{\Sigma x_i}{n}$	$\left(\frac{n+1}{2}\right)^{th}$	max. fq.
$\frac{\Sigma f_i x_i}{\Sigma f_i}$	$l + \left(\frac{\frac{N}{2} - cf}{f}\right)h$	$l + \left(\frac{f_1 - f_0}{2f_1 - f_0 - f_2}\right)h$

Quartile

$$Q_i = l + \left(\frac{\frac{iN}{4} - cf}{f}\right)h$$

$N = 1, 2, 3$

Decile

$$D_i = l + \left(\frac{\frac{iN}{10} - cf}{f}\right)h$$

$N = 1 - 9$

Percentile

$$P_i = l + \left(\frac{\frac{iN}{100} - cf}{f}\right)h$$

$N = 1 - 99$

Interpolation : estimating value b/w 2 values

Extrapolation : estimating value after values
(prediction)

Open-ended data : no upper limit

Q_2 : median value

$Q_2 : D_5 = P_{50}$

Group	Fq	Cf
0 - 10	8	8
10 - 20	12	20
20 - 30	20	40
30 - 40	32	72
40 - 50	30	102
50 - 60	28	130
60 - 70	12	142

$$\frac{146}{4} = 36.5$$

$$\frac{146}{10} = 14.6$$

$$\begin{array}{r} 14.6 \\ \times\ 7 \\ \hline 102.2 \end{array}$$

$$\frac{146}{100} = 1.46$$

$$\begin{array}{r} 1.46 \\ \times\ 85 \\ \hline 124.1 \end{array}$$

70 + 4 146

$$Q_1 = 20 + \frac{\frac{1 \cdot 146}{4} - 20}{20} \times 10$$

$$= 20 + \frac{36.5 - 20}{2}$$

$$= 20 + \frac{16.5}{2}$$

$$= \boxed{28.25}$$

$$D_7 = 50 + \frac{102.2 - 102}{28} \times 10$$

$$= 50 + \frac{2}{28}$$

$$= \boxed{50.071}$$

$$P_{85} = 50 + \frac{124.1 - 102}{28} \times 10$$

$$= 50 + \frac{221}{28}$$

$$= \boxed{57.892}$$

CI (e)	f_i	CI (inc)	x_i	$f_i x_i$	cf
15-19	4	14.5-19.5	17	68	4
20-24	20	19.5-24.5	22	440	24
25-29	38	24.5-29.5	27	1026	62
30-34	24	29.5-34.5	32	768	86
35-39	10	34.5-39.5	37	370	96
40-44	9	39.5-44.5	42	378	105
	105			3050	

$$\text{mean} = \frac{3050}{105} = \boxed{29.047} \qquad \frac{105}{2} = 52.5$$

$$\text{median} = 25 + \frac{52.5 - 24}{38} \times 5$$

$$= 25 + 3.75$$

$$= \boxed{28.75}$$

$$\text{mode} = 25 + \frac{38 - 20}{76 - 20 - 24} \times 5 \qquad \frac{18 \times 5}{32} = \frac{90}{32}$$

$$= \boxed{27.81}$$

Wages	fq	cf	UL
20-40	4	4	40
40-60	6	10	60
60-80	10	20	80
80-100	16	36	100
100-120	12	48	120
120-140	7	55	140
140-160	3	58	160

29

$Q_3 = \frac{58 \times 3}{4_2}$

$= \frac{87}{2}$

$= 43.5$

$$\frac{\alpha_1}{\alpha_2} = \frac{D_1}{D_2}$$

$$\frac{x - 100}{120 - 100} = \frac{43.5 - 36}{48 - 36}$$

$$\frac{x - 100}{20} = \frac{7.5}{12}$$

$$x - 100 = \frac{75 \times 2}{12} = \frac{150}{12} = 12.5$$

$$\boxed{x = 112.5}$$

$$\text{lower class boundary} + \frac{\text{how many in}}{\text{group total}} \times \text{class width}$$

$$P_{11} = \frac{58 \times 11}{100} = \underline{6.38}$$

$$\frac{\alpha_1}{\alpha_2} = \frac{D_1}{D_2}$$

$$\frac{x - 40}{60 - 40} = \frac{6.38 - 4}{10 - 4}$$

$$\frac{x - 40}{20} = \frac{2.38}{6}$$

$$x - 40 = \frac{23.8 \times 21}{63} \qquad 7.93$$

$$x = \boxed{47.93}$$

$$y = y_1 + \frac{(x - x_1)(y_2 - y_1)}{(x_2 - x_1)}$$

y: interpolated value

x: point to perform interpolath

x_1, y_1 : first coordinates

$x_2 y_2$: second coordinates

Marks	fq	cf
0-15	13	13
15-20	24	37
20-35	21	58
35-50	36	94
50-75	64	158
75-100	18	176

find 40-85

$$\frac{V - UL_U}{UL_C - UL_U} = \frac{V - CF_U}{CF_C - CF_U}$$

gradient of lines

$$\frac{40-35}{50-35} = \frac{x-58}{94-58}$$

$$\frac{85-75}{100-75} = \frac{x-158}{176-158}$$

$$\frac{8^1}{15\,3} = \frac{x-58}{36\,12}$$

$$\frac{2\,\cancel{10}}{5\,\cancel{25}} = \frac{x-158}{18}$$

$$x = 70$$

$$\frac{36}{5} = x - 158$$

$$165.2 - 70$$

$$= 95.2$$

$$7.2 = x - 158$$

$$x = 165.2$$

Salary (000's)	fq	cf	UL
0 - 15	17	17	15
15 - 20	28	45	20
20 - 30	87	132	30
30 - 50	122	254	50
50 - 65	89	343	65
65 - 80	37	380	80
> 80	23	403	

28k - 68k

$$28k: \frac{28-20}{30-20} = \frac{x-45}{132-45}$$

$$\frac{8}{10} = \frac{x-45}{87}$$

$$0.8 \times 87 = x - 45$$

$$69.6 = x - 45$$

$$\boxed{x = 114.6}$$

$$68k: \frac{68-65}{80-65} = \frac{x-343}{380-343}$$

$$\frac{3}{5.15} = \frac{x-343}{37\ 7.4}$$

$$\boxed{x = 350.4}$$

$$\begin{array}{r} {}^{4}\,{}^{9} \\ 350.4^{14} \\ -\ 114.6 \\ \hline 235.8 \end{array}$$

CI	f_i	cf	x_i	$f_i x_i$
100 - 200	83	83	150	12.450
200 - 300	91	174	250	22.750
300 - 500	104	278	400	41.600
500 - 700	72	350	600	43.200
700 - 1000	18	368	850	15.300
	368			135.300

$\frac{368}{2} = 184$

mean, median, mode, Q_3, D_7, P_{63}

1. mean $= \frac{135300}{368} = \boxed{367.66}$

2. median $= 300 + \frac{184 - 174}{104} \cdot 200$

$= 300 + \frac{2000}{104}$

$= \boxed{319.23}$

3. mode = $300 + \frac{104 - 91}{208 - 91 - 72} \cdot 200$

$208 - 163 = 45$

$91 + 72 = 163$

$= 300 + \frac{2600}{45}$

$= \boxed{357.77}$

4. $Q_3 : \frac{3 \times 368^{92}}{4_1} = 276$

$= 300 + \frac{276 - 174}{104} \cdot 200$

$= 300 + \frac{20400}{104}$

$= \boxed{496.153}$

5. $D_7 : \frac{7 \times 368}{10} = 257.6$

$= 300 + \frac{257.6 - 174}{104} \cdot 200$

$= 300 + \frac{16720}{104}$

$= \boxed{460.77}$

6. $P_{63} : \frac{63 \times 368}{100} = 231.84$

$$= 300 + \frac{231.84 - 174}{104} \cdot 200$$

$$= 300 + \frac{11,568}{104}$$

$$= \boxed{411.23}$$

Confidence interval	fq	cf
< 35	13	13
35 - 40	21	34
40 - 60	37	71
60 - 75	44	115
75 - 100	8	123

Marks b/lw 45 & 80 ?

$$\frac{45-40}{60-40} = \frac{x-34}{71-34} \qquad \frac{80-75}{100-75} = \frac{y-115}{123-115}$$

$$\frac{5'}{20'4} = \frac{x-34}{37} \qquad \frac{5'}{15'3} = \frac{y-115}{8}$$

$$\frac{37}{4} = x-34 \qquad \frac{8}{3} = y-115$$

$$x = 34+9.25 \qquad y = 115+2.67$$

$$\boxed{x = 43.25} \qquad \boxed{y = 117.67}$$

$$y - x = \underline{74.42}$$

Geometric mean:

$$(x_1 \times x_2 \times x_3 \times \ldots \times x_n)^{1/n}$$

Quartile deviation = $\frac{IQR}{2}$ (Semi Inter-quartile range)

$$QD = \frac{Q_3 - Q_1}{2}$$

Coefficient of quartile deviation (R) = $\frac{Q_3 - Q_1}{Q_3 + Q_1} \times 100$

(units) Absolute Measures of Deviation (ratios) Relative Measures of Deviatⁿ:

Coefficient of eg. mean

(compare data sets)

- Range
- Quartile Deviatⁿ
- Mean Deviatⁿ
- Standard Deviatⁿ

Coefficient of range (R): $\frac{Max - min}{Max + min} \times 100$

Residual: $x_i - \bar{x}$

Mean deviation: $\frac{\sum_{i=1}^{n} |x_i - \bar{x}|}{n}$

Degree of freedom: $n-1$

For grouped data: $\frac{\sum_{i=1}^{n} f_i |x_i - \bar{x}|}{n}$

Coefficient of mean deviation (R): $\frac{\text{mean deviatⁿ}}{\text{mean/median/mode}} \times 100$

Variance: $\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n}$ sum of squares of deviatⁿs from mean

Standard deviatⁿ: $\sigma = \sqrt{\text{Variance}}$ not affected by what you add/subtract in the values

Coefficient of variance (R): $\frac{sd}{mean} \times 100$

tens place / units place

Stem And Leaf Diagram

23	27	32	59	63	34	36	40
33	35	41	25	27	18	29	37
48	53	58	64	67	16	31	24
81	34	22	43	59	71	68	61
54	78	74	55	31	48	82	64

min value : 16
max value : 82

2	1 - 8, 6	2	1 - 6, 8
9	2 - 3, 7, 5, 7, 9, 4, 2	7	2 - 2, 3, 4, 5, 7, 7, 9
18	3 - 2, 4, 6, 3, 5, 7, 1, 4, 1	9	3 - 1, 1, 2, 3, 4, 4, 5, 6, 7
23	4 - 0, 1, 8, 3, 8	5	4 - 0, 1, 3, 8, 8
29	5 - 9, 3, 8, 9, 4, 5	6	5 - 3, 4, 5, 8, 9, 9
35	6 - 3, 4, 7, 8, 1, 4	6	6 - 1, 3, 4, 4, 7, 8
38	7 - 1, 8, 4	3	7 - 1, 4, 8
40	8 - 1, 2	2	8 - 1, 2

no mode

median = $\frac{41}{2} = 20.5^{th}$ obs = $\frac{41+43}{2} = \underline{\underline{42}}$

mean = 46.125

3 — 1. Title
2. Scales & Units
3. Naming Axes
} Diagram

$$Q_1 = \left(\frac{n+1}{4}\right)^{th} = \frac{41}{4}^{th} = 10.25^{th}$$

$$= 10^{th} + 0.25\,(11^{th} - 10^{th})$$

$$= 31 + 0.25(0)$$

$$= \underline{31}$$

$$Q_3 = 3\left(\frac{n+1}{4}\right)^{th} = 30.75^{th}$$

$$= 30^{th} + 0.75\,(31^{st} - 30^{th})$$

$$= 61 + 0.75\,(63 - 61)$$

$$= 61 + 0.75(2)$$

$$= 61 + 1.5$$

$$= \underline{62.5}$$

$$QD = \frac{Q_3 - Q_1}{2} = \frac{62.5 - 31}{2} = \frac{31.5}{2} = \underline{16.75}$$

$$\text{Coeff of } QD = \frac{Q_3 - Q_1}{Q_3 + Q_1} \times 100 = \frac{31.5}{93.5} \times 100 = \underline{33.68}$$

$$V(x) = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n}$$

$$= \frac{\sum_{i=1}^{n} (x_i^2 - 2x_i\bar{x} + \bar{x}^2)}{n}$$

$$= \frac{\sum_{i=1}^{n} x_i^2}{n} - \frac{2\bar{x}\sum_{i=1}^{n} x_i}{n} + \frac{\bar{x}\sum_{i=1}^{n} 1}{n}$$

$$= \frac{\Sigma x_i^2}{n} - 2\bar{x}^2 + \frac{\bar{x}^2(n)}{n} \#$$

$$= \frac{\Sigma x_i^2}{n} - 2\bar{x}^2 + \bar{x}^2$$

$$V(x) = \boxed{\frac{\Sigma x_i^2}{n} - \bar{x}^2}$$

$$\text{grouped: } \boxed{\frac{\Sigma f_i x_i^2}{n} - \bar{x}^2}$$

$$* \frac{\sum_{i=1}^{n} x_i}{n} = \bar{x} \qquad \sum_{i=1}^{n} 1 = n$$

Standard Deviation

i) Raw Data = (popn)

$$\sqrt{\frac{\Sigma x^2 - \frac{\Sigma(x)^2}{n}}{n}}$$

(use n-1 for sample)

ii) Discrete data =

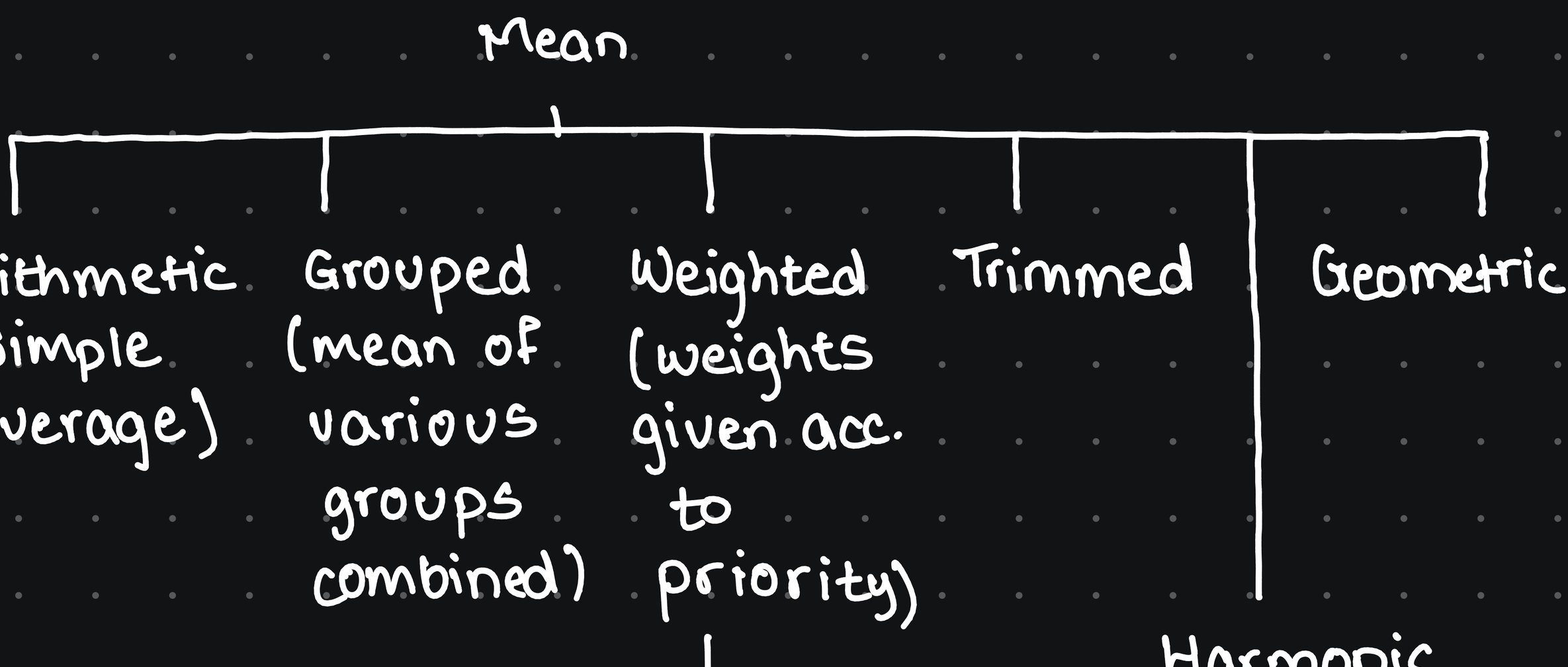
$$\sqrt{\frac{\Sigma f_i(x_i)^2}{n} - \left(\frac{\Sigma f_i x_i}{n}\right)^2}$$

iii) Continuous data =

$$c\sqrt{\frac{\Sigma f_i(d_i)^2}{n} - \left(\frac{\Sigma f_i d_i}{n}\right)^2}$$

$d = \frac{x - A}{c}$ — assumed mean

— c.i.



priority of observations is important
used when arithmetic mean does not show fair value

$$\text{mean} = \frac{\Sigma w_i x_i}{\Sigma w_i}$$ } used in index-number calculation

weights indicate significance of the number in the calculation

weights are treated as frequency while doing calculations

designed by researcher according to market trends

eg. Group 1 | Group 2 | Group 3

$N_1 = 20$ | $N_2 = 30$ | $N_3 = 25$

$\bar{x}_1 = 104$ | $\bar{x}_2 = 123$ | $\bar{x}_3 = 116$

$$\text{grouped mean} = \frac{N_1\bar{x}_1 + N_2\bar{x}_2 + N_3\bar{x}_3}{N_1 + N_2 + N_3}$$

$$= \frac{2080 + 3690 + 2900}{75}$$

$$= \underline{115.6}$$

eg.	x_i	weights(w_i)	$w_i x_i$
	1800		
	12000	34	4.08.000
	56000		
	120000	4	4.80.000
	9200		
	6000		
	23000	57	13.11.000

eg.

Commodity	Consumption for 100 ppl in quintals (x_i)	w_i	$w_i x_i$
Rice	143	56	8.008
Wheat	231	67	15.477
Oil	93	41	3.813
Lentils	32	23	736
Sprouts	21	6	126
Rawa (Semolina)	7	2	14
		$\Sigma = 195$	$\Sigma = 28.174$

$$\text{weighted mean} = \frac{\Sigma w_i x_i}{\Sigma w_i}$$

$$= \frac{28174}{195}$$

$$= 144.48$$

eg. Program	Casting hours	revenue (millions)
football match	132	567
Cricket	89	342
movies	45	14
Cartoons	164	752
	Σ= 430	

find weighted mean. taking weight as percentage of casting hours —— refer stats sheet 4.

* SET B *

1. (d) 20

$$\text{coeff of range} = \frac{L-S}{L+S} \times 100$$

$$= \frac{90-60}{90+60} \times 100$$

$$= \frac{30}{150} \times 100$$

$$= 20$$

2.

3. (c) 72.46

$$\frac{59.5-9.5}{59.5+9.5} \times 100$$

$$= \frac{50}{69} \times 100$$

$$= 72.46$$

4. (c) -6

5. (c) 1.44

x_i	$\lvert x - x_i \rvert$
5	0.2
8	2.8
6	0.8
3	2.2
4	1.2
$\Sigma = 26$	$\Sigma = 7.2$

$$\bar{x} = \frac{26}{5} = 5.2$$

$$MD = \frac{\Sigma |x_i - x|}{n}$$

$$MD = \frac{7.2}{5} = 1.44$$

6. (c) 35

x_i	$f_i \lvert x_i - \bar{x} \rvert$	f_i	$f_i x_i$	$\lvert x_i - \bar{x} \rvert$
50	35	7	350	5
60	35	7	420	5
		14	770	10

$$\bar{x} = \frac{770}{14} = 55$$

$$MD = \frac{70}{2} = 35$$

7. (c) 400/9

$$\Sigma 1\text{-}9 = \frac{9 \times 10}{2} = 45$$

$$\Sigma = \frac{n \times (n+1)}{2}$$

$$\bar{x} = \frac{45}{9} = 5$$

$$\tilde{x} = \frac{n+1}{2}$$

x_i	$\lvert x_i - \bar{x} \rvert$
1	4
2	3
3	2
4	1
5	0
6	1
7	2
8	3
9	4

$\Sigma = 20$

$$MD = \frac{20}{9}$$

$$CoMD = \frac{\frac{20}{9}}{5} \times 100 = \frac{20}{9 \times 5} \times 100 = \frac{400}{9}$$

Geometric Mean

$$GM = \sqrt[n]{x_1 x_2 \dots x_n}$$

situations where data values are unitless & depend on each other

(eg. index nos.
roi

interdependent

AM does not do justice

find r. change of figures that build/compound on each other

yr	ror
1	0.1
2	0.22
3	0.06
4	-0.05
5	0.2

$A = P(1+i)^n$

compound int is another way of finding GM!!

$A = (1+i_1)(1+i_2)(1+i_3)\dots$

& $P = 1$

$(1+i) = \sqrt[n]{x_1 x_2 x_3 \dots x_n}$

yr	ror	1 + ror
1	0·10	1·10
2	0·22	1·22
3	0·06	1·06
4	-0·05	0·95
5	0·20	1·20

product = 1·621673

$\sqrt[5]{1.621673} = 1.10152069 = 1+i$

$\therefore i = 0.102$

$i = \frac{r}{100}$

$\therefore r = 10.2\%$

compounded annual growth rate (CAGR)

Geometric mean of grouped data

1) Discrete (for continuous, take mid-value as x_i)

$$G = (x_1^{f_1} \times x_2^{f_2} \times \dots \times x_n^{f_n})^{1/N}$$

eg.

x_i	f_i
2	2
4	3
8	3
16	2

$\Sigma = 10$

$$G = (2^2 \times 4^3 \times 8^3 \times 16^2)^{1/10}$$

$$= (2)^{2.5}$$

$$= 4\sqrt{2}$$

$$= 5.66$$

$\sqrt{2} = 1.41$

$\sqrt{3} = 1.712$

$\sqrt{10} \approx pi$

investopedia

i) Logarithm of G for a set of obs. is the AM of the logarithm of the obs.

$$\log G = \frac{1}{n} \Sigma \log x$$

ii) if all the obs. assumed by a variable are constants. say $k > 0$. then the GM of the obs. is also k

iii) GM of the product of two variables is the product of their GMs.

$$z = xy$$

$$GM(z) = GM(x) \times GM(y)$$

$$z = \frac{x}{y}$$

$$GM(z) = \frac{GM(x)}{GM(y)}$$

* $R_y = |m| \times R_x$
(range y) (range x)

linearity property of measures of dispersn

$2x + 3y = 10$

$R_x = 15$

$R_y = ?$

$2x + 3y = 10$

$y = \frac{10}{3} - \frac{2}{3}x$ $(y = mx + c)$

shift of origin & scale (c)

$R_y = |m| R_x = \frac{2}{3} \times 15 = \underline{\underline{10}}$ (a)

Q. AM = 6.5 Find nos

GM = 6

$$\Rightarrow \quad 6 = \sqrt{xy}$$

$$36 = xy$$

$$\frac{x+y}{2} = 6.5$$

$$x+y = 13$$

(ii) 9 & 4

Q. $4x + 3y + 11 = 0$

$MD_x = 5.4$

$MD_y = ?$

$$4x + 3y = -11$$

$$y = \frac{-11}{3} - \frac{4x}{3}$$

$$m = -\frac{4}{3} \qquad |m| = \frac{4}{3}$$

$$MD_y = |m| MD_x = \frac{4}{\cancel{3}_{1}} \overset{1.8}{\cancel{(5.4)}} = \underline{7.2}$$

change of scale :

In the y=mx+c scale, if we change the origin constant, it means shifting the entire graph up or down without changing its slope. This change does not affect the linearity of the graph.

$y = mx + c$

↓ add

change of origin

In the y=mx+c scale, if we change the scale constant (m), it means stretching or compressing the graph vertically without changing its slope. This change does not affect the linearity of the graph.

⇓ multiply

change of scale: multiply

change of constant: add

linearity not affected by change in constants

y: mate

change of origin: add

constant

linearity not affected by change in constants

1. If all obs. assumed by a variable are constant (equal), then SD is zero

   Also applies to range & mean deviat^n

2. SD unaffected due to change of origin but affected in same ratio due to change of scale

$$y = mx + c$$

$$S_y = |m| S_x$$

3. obs: n_1 & n_2
   AM : $\bar{x}_1$ & $\bar{x}_2$ of 2 Groups
   SD : s_1 & s_2

Combined SD

$$S = \sqrt{\frac{n_1 s_1^2 + n_2 s_2^2 + n_1 d_1^2 + n_2 d_2^2}{n_1 + n_2}}$$

$$= \sqrt{\frac{n_1(\sigma_1^2 + d_1^2) + n_2(\sigma_2^2 + d_2^2)}{n_1 + n_2}}$$

$$d_1 = \bar{x}_1 - \bar{x}$$

$$d_2 = \bar{x}_2 - \bar{x}$$

$$\bar{x} = \frac{n_1 \bar{x}_1 + n_2 \bar{x}_2}{n_1 + n_2}$$

$3x+2y+10=0$

Relath bllw $R_x + R_y$?

$3x+2y=-10$

$y = -5 - \frac{3x}{2}$

$R_y = \frac{3}{2} R_x$

$\boxed{2R_y = 3R_x}$

Q. $2x+3y-7=0$

$\bar{x}_x = 1$

$MD_{x(mean)} = 0.3$

$COMD_{x(mean)} = ?$

$2x+3y=7$

$y = \frac{7}{3} - \frac{2x}{3} \therefore |m| = \frac{2}{3}$

$\bar{x}_y = \frac{2}{3}$. $MD_{y(mean)} = 0.2$

$COMD_y = 0.2 \times \frac{3}{2} \times 100 = \underline{\underline{30}}$

$COMD_x = \frac{0.3 \times 100}{1}$

$= 30$

$COMD_y = \frac{2}{3} \times 30$

$= 20$

if $2x + 3y - 7 = 0$,

$2\bar{x} + 3\bar{y} - 7 = 0$

($\bar{x}$ & $\bar{y}$ are AMs)

$\therefore 2(1) + 3\bar{y} = 7$

$3\bar{y} = 5$

$$\boxed{\bar{y} = \frac{5}{3}}$$

$MD_{y(mean)} = 0.2$

$COMD_{y(mean)} = 0.2 \times \frac{3}{5} \times 100$

$= \frac{60}{5}$

$= 12$

$$\boxed{(b)\ 12}$$

Q. $\bar{x}_x = a$

$SD_x = b$

SD of $\frac{x-a}{b}$ = ?

$y = mx + c$

$m = \frac{1}{b}$

$\Rightarrow \quad y = \frac{x-a}{b} = \frac{x}{b} - \frac{a}{b}$

$SD_y = \frac{1}{b}(SD_x)$

$= \frac{1}{b}(b)$

$= 1$

Q. $SD_x = 3$

σ of $(5-2x)$ = ?

$\Rightarrow \quad y = 5 - 2x$

$m = 2$

$\sigma = \sigma^2$

$SD_y = 2(3) = 6$

$\therefore \sigma_y = 36$

Q. $2x + 3y + 4 = 0$

$SD_x = 6$

$SD_y = ?$

$2x + 3y = -4$

$y = \frac{-4}{3} - \frac{2x}{3}$

$m = \frac{2}{3}$

$SD_y = \frac{2}{3}(6) = \underline{4}$

Q. G_1 G_2

$n_1 = 100$ $n_2 = 150$

$\bar{x}_1 = 31$ $\bar{x}_2 = 37$

$SD_{x_1} = 4.3$ $SD_{x_2} = 5.3$

$S_1^2 = 18.49$ $S_2^2 =$

$$S = \sqrt{\frac{n_1 s_1^2 + n_2 s_2^2 + n_1 d_1^2 + n_2 d_2^2}{n_1 + n_2}}$$

$$d_1 = \bar{x}_1 - \bar{x}$$

$$d_2 = \bar{x}_2 - \bar{x}$$

$$\bar{x} = \frac{n_1 x_1 + n_2 x_2}{n_1 + n_2}$$

$$\bar{x} = \frac{3100 + 5550}{250} = 34.6$$

$d_1^2 = 12.96$ $d_2^2 = 5.76$

$d_1 = -3.6$ $d_2 = 2.4$

$$S = \sqrt{\frac{1849 + 4213.5 + 1296 + 864}{250}}$$

$$= \underline{5.73}$$

Year	Popn %	i (r/100)	1 + i
1994	5.3%	0.053	1.053
1995	6.2%	0.062	1.062
1996	7.1%	0.071	1.071
1997	8.1%	0.081	1.081
1998	-1.5%	-0.015	0.985
1999	-2%	-0.02	0.980
2000	3.1%	0.031	1.031

avg %

product of $(1+i) = 1.288513516$

7th root $= 1.0369$

$1 + i = 1.0369$

$i = 0.0369$

$i = \frac{r}{100}$

$\therefore \underline{r = 3.69\%}$

Harmonic mean

reciprocal of AM of reciprocals of obs.

4, 7, 10, 13...

$HP = \frac{1}{4}, \frac{1}{7}, \frac{1}{10}$ ....

$$HM = \frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}}$$

none of the obs. should be 0.

ungrouped.

Used in averaging of ratios. (fractions)

both variables vary. (diff to keep one thing constant)

$AM \geq GM \geq HM$

*adv

rigid

based on all obs.

fluctuat^n doesn't affect

more weight to smaller items

*disadv

not easily understandable

diff. to compute

2 $\frac{1}{2}$

3 $\frac{1}{3}$

$GM^2 = AM \times HM$

$GM = \sqrt{AM \times HM}$

Applicat^n :

R — Rate Metric — rate changes acc. to many factors

A — Async Time Intervals — eg. canteen
asynchronisat^n of time intervals

T — Too low data values are of interest — small values are significant

E — Event data.. not sampled data — observat^al. record as event occurs

Grouped

$$HM = \frac{n}{\sum_{i=1}^{n} \left[\frac{f_i}{x_i} \right]}$$

Combined HM

$$HM = \frac{n_1 + n_2}{\frac{n_1}{H_1} + \frac{n_2}{H_2}}$$

Q. HM of 4, 6, 10

$$HM = \frac{n}{\Sigma \frac{1}{x_i}}$$

$$\frac{1}{4} + \frac{1}{6} + \frac{1}{10} = \frac{15+10+6}{60} = \frac{31}{60}$$

$$HM = \frac{3}{\frac{31}{60}} = 3 \times \frac{60}{31} = \frac{180}{31} = \underline{\underline{5.81}}$$

Q. HM

x_i	f_i	f_i/x_i
2	2	$\frac{2}{2} = 1$
4	3	$\frac{3}{4} = 0.75$
8	3	$\frac{3}{8} = 0.375$
16	2	$\frac{2}{16} = \frac{1}{8} = 0.125$
	10	

$$HM = \frac{10}{2.25} = \underline{\underline{4.44}}$$

Q. AM. GM. HM For 6, 8, 12, 36

$$AM = \frac{6+8+12+36}{4} = \frac{62}{4} = \underline{\underline{15.5}}$$

$$GM = \sqrt[4]{6 \cdot 8 \cdot 12 \cdot 36} = \sqrt[4]{2^4 6^4} = \underline{\underline{12}}$$

$$HM = \frac{4}{\frac{1}{6} + \frac{1}{8} + \frac{1}{12} + \frac{1}{36}}$$

$$= \frac{4}{0.167 + 0.125 + 0.083 + 0.028}$$

$$= \frac{4}{0.403}$$

$$= \underline{\underline{9.93}}$$

$$\boxed{GM^2 = AM \cdot HM}$$

$$\text{Weighted GM} = \text{Antilog}\left(\frac{\Sigma w_i \log x_i}{\Sigma w_i}\right)$$

$$\text{Weighted HM} = \frac{\Sigma w_i}{\Sigma\left(\frac{w_i}{x_i}\right)}$$

$$\text{Weighted AM} = \frac{\Sigma w_i x_i}{\Sigma w_i}$$

$$\text{Mode} \simeq 3\,\text{median} - 2\,\text{mean}$$

Q. Find weighted AM & HM of first 'n' natural nos., the weights being equal to the square of corresponding nos.

$\Rightarrow \quad w_i = (n_i)^2$

$$AM = \frac{\Sigma w_i x_i}{\Sigma w_i} = \frac{\Sigma n^2 \cdot n}{\Sigma n^2} = \frac{\Sigma n^3}{\Sigma n^2}$$

$$\Sigma n^2 = \frac{n(n+1)(2n+1)}{6}$$

$$\Sigma n^3 = \left(\frac{n(n+1)}{2}\right)^2$$

$$\therefore AM = \frac{\frac{\cancel{n}(\cancel{n}+1)}{\cancel{2}} \cdot \frac{n(n+1)}{2}}{\frac{\cancel{n}(n+1)(2n+1)}{\cancel{6}\ 3}}$$

$$\boxed{AM = \frac{3n(n+1)}{2n+1}}$$

$$HM = \frac{\Sigma w_i}{\Sigma\left(\frac{w_i}{x_i}\right)} = \frac{\Sigma n^2}{\Sigma n} = \frac{\frac{\cancel{n}(n+1)(2n+1)}{\cancel{6}\ 3}}{\frac{\cancel{n}(n+1)}{\cancel{2}}}$$

$$\frac{w_i}{x_i} = \frac{n^2}{n} = n$$

$$\boxed{HM = \frac{2n+1}{3}}$$

8.

Year	Dep. value	i	1-i	
2013	25%	0.25	0.75	
2014	10%	0.10		} 0.90
2015	10%	0.10		
2016	10%	0.10		
2017 - 2021	2%	0.02	0.98	

original cost = 50 L
find cost @ 2021 end.

$$\Rightarrow \quad 0.75 \times (0.90)^3 \times (0.98)^5 = (1-i)^9$$

$$0.4942186957 = (1-i)^9$$

9th root $\simeq 0.92467905$

$$1 - i = 0.93$$

$$i = 0.07$$

$r = 7\%$ (avg rate)

$$A = 50.00.000 \times 0.4942$$

$A \simeq$ 24,71,000 approx.

$$A = P(1+i_1)(1+i_2)^n \ldots\ldots$$

Q. A car travelled the distance with 4 speeds:
50 mph, 20mph, 40mph, 25 mph.
Find average speed

⇓

Harmonic mean always.

$$\Rightarrow HM = \frac{4}{\frac{1}{50} + \frac{1}{20} + \frac{1}{40} + \frac{1}{25}}$$

$$= \frac{4}{0.02 + 0.05 + 0.025 + 0.04}$$

$$= \frac{4}{0.135}$$

$$= 29.63 \text{ mph}$$

8. $n_1 = 4$ $GM_1 = 47$

$n_2 = 6$ $GM_2 = 40$

$GM\,(all\ 10) = ?$

$$\Rightarrow \sqrt[10]{47^4 * 40^6}$$

$$= \sqrt[10]{4879681 \times 4096000000}$$

$$= \sqrt[10]{19987173376000000}$$

$$= \underline{42.67}$$

Combined GM

$$= \sqrt[n_1+n_2]{G_1^{n_1} \cdot G_2^{n_2}}$$

$$\log G = \frac{1}{n_1+n_2}\left[n_1 \log g_1 + n_2 \log g_2\right]$$

Q. GM = 3.63

HM = 3.27

AM = 4

n = 3

Find obs.

trial & error method

$\frac{a+b+c}{3} = 4$

2.4.6

$a+b+c = 12$

$\sqrt[3]{abc} = 3.63$

$abc \simeq 47.832147 \approx 48$

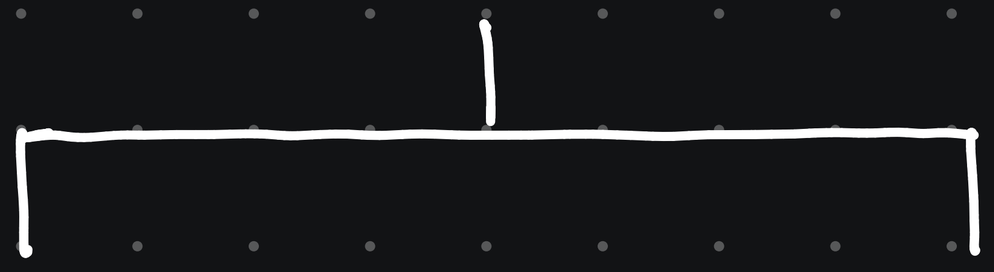
$\frac{3}{\frac{ab+bc+ca}{abc}} = 3.27$

$\frac{3(47.832147)}{ab+bc+ca} = 3.27$

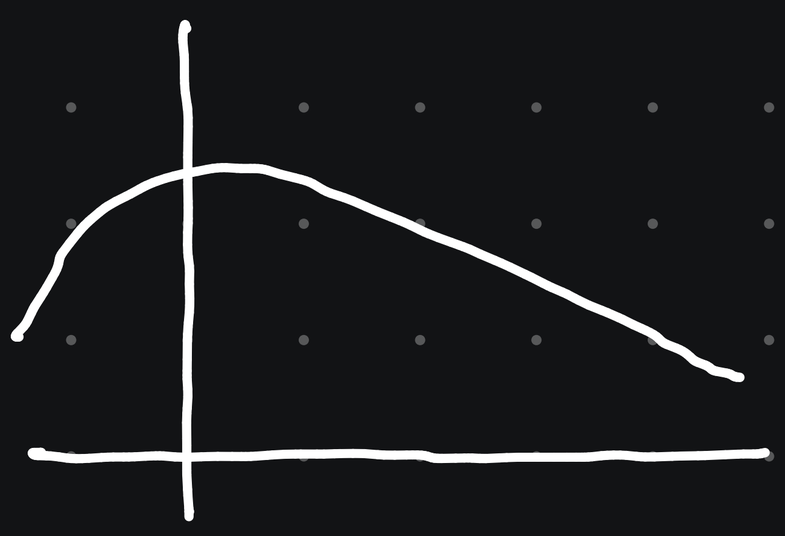
$ab+bc+ca \simeq 43.883$

Skewness — measure of asymmetry

tendency of data not showing symmetry
(going away from symmetry)



Positive
(right-tailed)



Negative
(left-tailed)

Skewness measures (coefficients)

Bowley's coeff
(open-ended data)

$-1 \leq Sk_b \leq 1$

$$\frac{Q_1 + Q_3 - 2Q_2}{Q_3 - Q_1} = Sk_p$$

entirely based on
quartiles
↓
$Q_1 - Q_3$ give more info
abt middle 50% part

Karl-Pearson's coeff

$$\frac{\text{mean} - \text{mode}}{SD}$$

$-1 \leq Sk_p \leq 1$

v.v. sensitive

more precise value

Q. Find S_b & Skp. Interpret.

CI	fi	cf		xi	fixi
5-10	4	4		7.5	30
10-20	13	17		15	195
20-30	29	46	Q_1	25	725
30-40	53	99	Q_2	35	1855
40-60	37	136	Q_3	50	1850
60-80	18	154		70	1260
80-100	9	163		90	810
	163				6725

① Bowley's = $\frac{Q_1 + Q_3 - 2Q_2}{Q_3 - Q_1}$

$\frac{163}{4} = 40.75$

$\times 2 = 81.5$

$\times 3 = 122.25$

$Q_1 : 20 + \frac{40.75 - 17}{29} \cdot 10 = 28.20$

$Q_2 = 30 + \frac{81.5 - 46}{53} \cdot 10 = 36.70$

$Q_3 = 40 + \frac{122.25 - 99}{37} \cdot 20 = 52.60$

$S_b = \frac{7.4}{24.4} = \boxed{0.303}$

② $KP = \frac{\text{mean} - \text{mode}}{SD}$

$\bar{x} = \frac{6725}{163} = 41.26$

$\text{mode} = 30 + \left(\frac{53-29}{2\cdot53-29-37}\right)10$

$= 30 + \frac{240}{40}$

$= 36$

f_i	x_i	x_i^2	$f_i(x_i)^2$
4	7.5	56.25	225
13	15	225	2925
29	25	625	18125
53	35	1225	64925
37	50	2500	92500
18	70	4900	88200
9	90	8100	72900
163			339800

$\sigma^2 = \frac{\Sigma f_i(x_i)^2}{n} - (\bar{x})^2 = 2084.66 - 1702.40$

$= 382.26$

$\sigma = 19.56$

$KP = \frac{41.26-36}{19.56} = \frac{5.26}{19.56} \simeq \boxed{0.27}$

Interpretation

positively skewed (more data after median value (Q_2))

right tailed distribut^n

mode < median < mean

$$Q_3 - Q_2 > Q_2 - Q_1$$

* Characteristics

Negatively skewed (more data before median value (Q_2))

left tailed distribut^n

mode > median > mean

$$Q_3 - Q_2 < Q_2 - Q_1$$

* mean - mode = 3(mean - median)

Combined S.D.= 11.01651655

→ Find:

Coeff. of variance

quartile deviath

Coeff. of QD

Skewness Sb

Skewness Skp

total money spent on wages (Interpret)

H.W. ⇒ created discrepancies

The following facts were gathered from a firm before and after an industrial dispute

	before	after
mean wages	185	190
median wages	182	180
mode	176	160
quartiles q1 & q3	175 AND 192	175 AND 195
S.D.	13	19
NO. OF EMPLOYEES	600	550

TO COMPARE THE POSITION OF THE FIRM BEFORE AND AFTER THE DISPUTE BY MAKIG USE OF THE DATA AS FULLY AS POS

DISPERSION -INTRO | dispersion-div B-problem | M.D.(GROUPED) | div-A | skewness | COMBINED s.d. | moments

25°C Partly sunny ENG 11:30 25-09-2023

technically applied to find skewness & kurtosis

Moments

deviat^n's from a fixed pt. raised to power

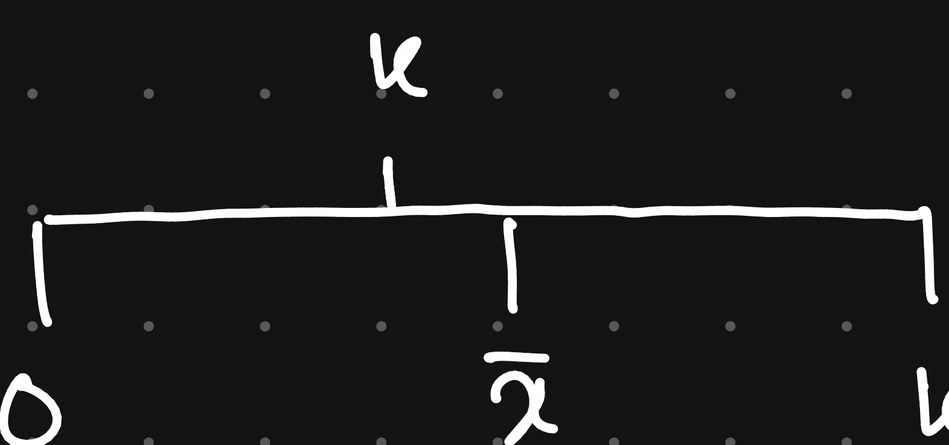
fix one reference point & find deviation of data

$x_1 \quad x_2 \quad \ldots \quad x_n \rightarrow$ data

fix point = k

deviations → $x_1 - k \quad x_2 - k \ldots \ldots \quad x_n - k$

$$\frac{\Sigma (x_i - k)^p}{n} = p^{th} \text{ moment}$$



raw moment — Central moment

Raw moment: $m'_r / \mu'_r = \frac{\sum_{i=1}^{n} x_i^{\,r}}{n}$

$m'_1 = \frac{\Sigma x_i}{n}$ (mean) $\qquad m'_3 = \frac{\Sigma x_i^3}{n}$

$m'_2 = \frac{\Sigma x_i^2}{n}$ $\qquad m'_4 = \frac{\Sigma x_i^4}{n}$

('avg sum of squares') → Avg s.s.

Central moment : $m_r / \mu_r = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^r}{n}$

GAP \

$m_1 = \frac{\Sigma(x_i - \bar{x})}{n} = 0$

$m_2 = \frac{\Sigma(x_i - \bar{x})^2}{n}$

$m_3 = \frac{\Sigma(x_i - \bar{x})^3}{n}$

0 for symmetric data

$m_4 = \frac{\Sigma(x_i - \bar{x})^4}{n}$

variance

Q.

x_i	x^2	x^3	x^4	$(x_i-\bar{x})$	$(x_i-\bar{x})^2$	$(x_i-\bar{x})^3$	$(x_i-\bar{x})^4$
5	25	125	625	2	4	8	16
3	9	27	81	0	0	0	0
1	1	1	1	-2	4	-8	16
4	16	64	256	1	1	1	1
2	4	8	16	-1	1	-1	1
15	55	225		0	10	0	34

$$\bar{x} = \frac{15}{5} = 3$$

Raw

$$m'_1 = \frac{15}{5} = 3$$

$$m'_2 = \frac{55}{5} = 11$$

$$m'_3 = \frac{225}{5} = 45$$

$$m'_4 = \frac{979}{5} = 195.8$$

Central

$$m_1 = 0$$

$$m_2 = \frac{10}{5} = 2$$

$$m_3 = 0$$

$$m_4 = \frac{34}{5} = 6.8$$

$$m_2 = \mu(x) = \frac{\Sigma x_i^2}{n} - (\bar{x})^2 = m'_2 - (m'_1)^2$$ ✓

$$m_3 = m'_3 - 3m'_1 m'_2 + 2(m'_1)^3$$ ✓

$$m_4 = m'_4 - 4m'_3 m'_1 + 6m'_2 (m'_1)^2 - 3(m'_1)^4$$ ✓

Binomial theorem

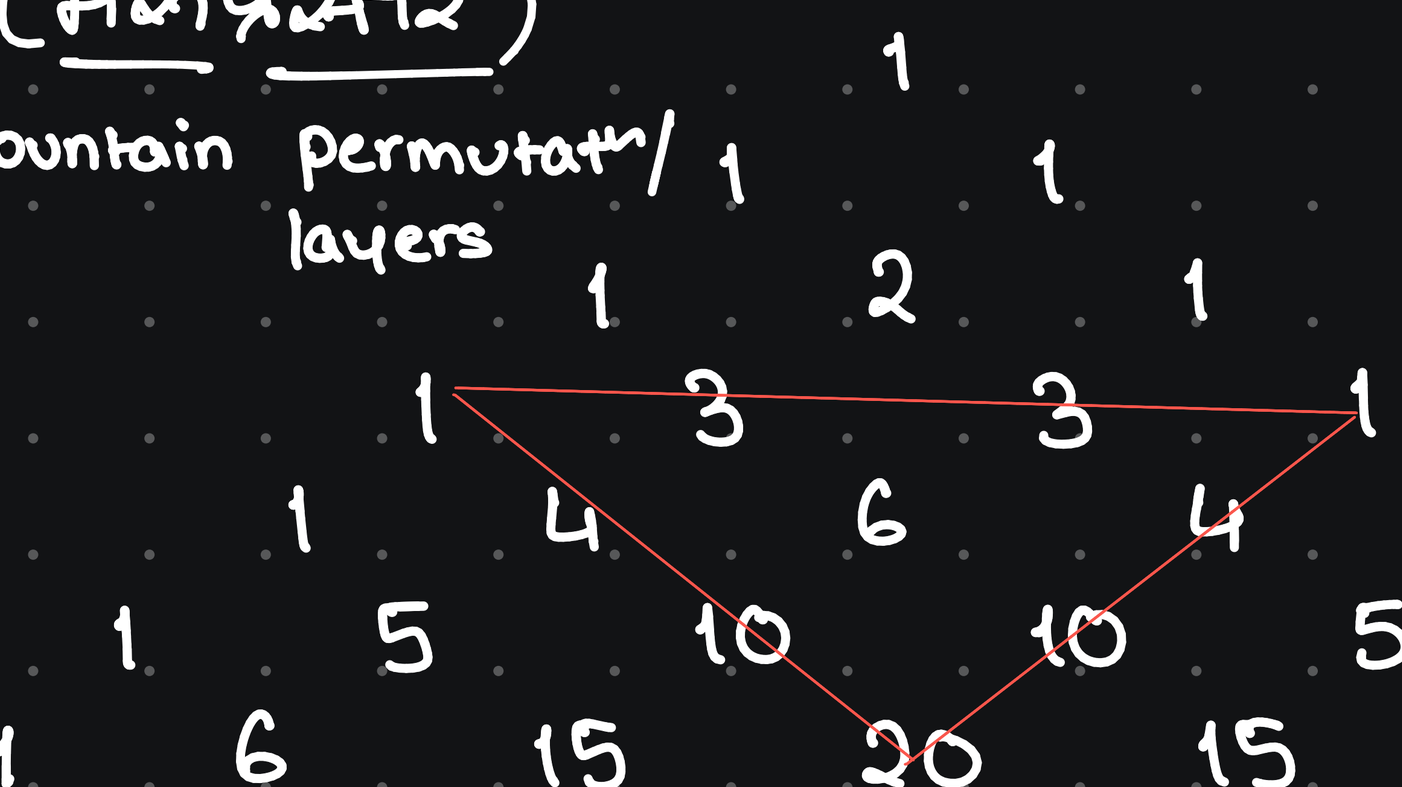
$$(x+y)^n = nC_0\, x^n y^0 + nC_1\, x^{n-1} y^1 + nC_2\, x^{n-2} y^2 + \ldots\ldots + nC_n\, x^{n-n} y^n$$

(Pascal - Hemchandra triangle)

Pascal's triangle for coefficients: — Given by पिंगलाचार्य

(मेरुप्रस्तार) — Δ of diff. combinatⁿ coeff.

mountain / permutatⁿ / layers



sum of squares of base

$(x+y)^0$
$(x+y)^1$
$(x+y)^2$
$(x+y)^3$
$(x+y)^4$
$(x+y)^5$
$(x+y)^6$ ....

eg. $(x+y)^5 = x^5 + 5x^4y + 10x^3y^2 + 10x^2y^3 + 5xy^4 + y^5$

hockey-stick property

$$m_2 = \frac{\Sigma (x_i - \bar{x})^2}{n}$$

$$= \frac{\Sigma (x_i^2 - 2x_i\bar{x} + \bar{x}^2)}{n}$$

$$= \frac{\Sigma x_i^2}{n} - \frac{2\bar{x}\Sigma x_i}{n} + \frac{\bar{x}^2 \Sigma 1}{n}$$

$$= \frac{\Sigma x_i^2}{n} - 2\bar{x}^2 + \bar{x}^2$$

$$= \frac{\Sigma x_i^2}{n} - \bar{x}^2$$

$$\boxed{m_2 = m'_2 - m'^2_1}$$

$$m_3 = \frac{\Sigma (x_i - \bar{x})^3}{n}$$

$$= \frac{\Sigma (x_i^3 - 3x_i^2\bar{x} + 3x_i\bar{x}^2 - \bar{x}^3)}{n}$$

$$= \frac{\Sigma x_i^3}{n} - 3\bar{x}\frac{\Sigma x_i^2}{n} + 3\bar{x}^2\frac{\Sigma x_i}{n} - \bar{x}^3\frac{\Sigma 1}{n}$$

$$= \frac{\Sigma x_i^3}{n} - 3\bar{x}\frac{\Sigma x_i^2}{n} + 3\bar{x}^3 - \bar{x}^3$$

$$= \frac{\Sigma x_i^3}{n} - 3\bar{x}\frac{\Sigma x_i^2}{n} + 2\bar{x}^3$$

$$\boxed{m_3 = m'_3 - 3m'_1 m'_2 + 2m'^3_1}$$

$$m_4 = \frac{\Sigma (x_i - \bar{x})^4}{n}$$

$$= \frac{\Sigma (x_i^4 - 4 x_i^3 \bar{x} + 6 x_i^2 \bar{x}^2 - 4 x_i^1 \bar{x}^3 + \bar{x}^4)}{n}$$

$$= \frac{\Sigma x_i^4}{n} - 4\bar{x} \frac{\Sigma x_i^3}{n} + 6\bar{x}^2 \frac{\Sigma x_i^2}{n} - 4\bar{x}^3 \frac{\Sigma x_i}{n} + \bar{x}^4 \frac{\Sigma 1}{n}$$

$$= \frac{\Sigma x_i^4}{n} - 4\bar{x} \frac{\Sigma x_i^3}{n} + 6\bar{x}^2 \frac{\Sigma x_i^2}{n} - 4\bar{x}^4 + \bar{x}^4$$

$$= \frac{\Sigma x_i^4}{n} - 4\bar{x} \frac{\Sigma x_i^3}{n} + 6\bar{x}^2 \frac{\Sigma x_i^2}{n} - 3\bar{x}^4$$

$$m_4 = m'_4 - 4m'_1 m'_3 + 6m'^2_1 m'_2 - 3m'^4_1$$

$$\boxed{m_4 = m'_4 - 4m'_1 m'_3 + 6m'_2 (m'_1)^2 - 3m'^4_1}$$

Skewness coefficients

Beta & Gamma coefficients

$$\beta_1 = \frac{m_3^2}{m_2^3}$$

$$\gamma_1 = \sqrt{\beta_1} = \text{Skewness}$$

if $m_3 < 0$, then $\gamma_1 < 0$

$$\beta_2 = \frac{m_4}{m_2^2}$$

$$\gamma_2 = \beta_2 - 3 = \text{kurtosis}$$

height of standard distribution = 3 units

⇓

mesokurtic

$\beta_2 > 3$	leptokurtic	$\gamma_2 > 0$
$\beta_2 = 3$	mesokurtic	$\gamma_2 = 0$
$\beta_2 < 3$	platykurtic	$\gamma_2 < 0$

No.	Data	$\mu = x - 13$	μ_i^2	μ_i^3	μ_i^4
1	12	-1	1.	-1	1
2	13	0	0	0	0
3	16	~~3~~	9.	~~27~~	81
4	11	~~-2~~	4.	~~-8~~	16
5	9	-4	16.	-64	256
6	15	~~2~~	4.	~~8~~	16
7	19	6	36.	216	1296
8	10	~~-3~~	9	~~-27~~	81
		1	79	151	1747

1) find first 4 Raw moments

2) Using relatn b/w Raw & Central moments. find central moments

3) Find Skewness & Kurtosis coeffs.

4) Interpret result

$\Rightarrow \underline{m'_1} = \frac{\Sigma x_i}{n} = \frac{1}{8} = \underline{0.125}$

$\underline{m'_2} = \frac{\Sigma x_i^2}{n} = \frac{79}{8} = \underline{9.875}$

$\underline{m'_3} = \frac{\Sigma x_i^3}{n} = \frac{151}{8} = \underline{18.875}$

$\underline{m'_4} = \Sigma x_i^4 / n = 1747/8 = \underline{218.375}$

} moments around 13

$$\underline{m_1} = \underline{0}$$

$$\underline{m_2} = m'_2 - m'^{2}_1$$

$$= 9.875 - (0.125)^2$$

$$= 9.859375$$

$$\underline{m_3} = m'_3 - 3m'_1 m'_2 + 2m'^{3}_1$$

$$= 18.875 - 3(0.125)(9.875) + 2(0.125)^3$$

$$= 18.875 - 3.703125 + 0.00390625$$

$$= \underline{15.17578125}$$

$$\underline{m_4} = m'_4 - 4m'_1 m'_3 + 6m'_2 (m'_1)^2 - 3m'^{4}_1$$

$$= 218.375 - 4(0.125)(18.875) + 6(9.875)(0.125)^2$$

$$- 3(0.125)^4$$

$$= 218.375 - 9.4375 + 0.92578125 - 0.000732421875$$

$$= \underline{209.8625488281}$$

$$\beta_1 = \frac{m_3^2}{m_2^3} = \frac{230.4324}{958.585256} = 0.2403880078$$

$$\gamma_1 \text{ (skewness)} = 0.4902937974$$

$$= 0.49 \text{ (positively skewed)}$$

$$\beta_2 = \frac{m_4}{m_2^2} = \frac{209.86}{97.2196} = 2.158618221$$

$$\gamma_2 \text{ (kurtosis)} = \beta_2 - 3$$

$$= 2.16 - 3$$

$$= -0.84 \text{ (platykurtic)}$$

	A	B	C	D	E	F	G	H	I	J
1		xi	xi^2	xi^3	xi^4	xi-x	(xi-x)^2	(xi-x)^3	(xi-x)^4	
2		12	144	1728	20736	-1.125	1.265625	-1.42382813	1.601806641	
3		13	169	2197	28561	-0.125	0.015625	-0.00195313	0.000244141	
4		16	256	4096	65536	2.875	8.265625	23.76367188	68.32055664	
5		11	121	1331	14641	-2.125	4.515625	-9.59570313	20.39086914	
6		9	81	729	6561	-4.125	17.01563	-70.1894531	289.5314941	
7		15	225	3375	50625	1.875	3.515625	6.591796875	12.35961914	
8		19	361	6859	130321	5.875	34.51563	202.7792969	1191.328369	
9	8	10	100	1000	10000	-3.125	9.765625	-30.5175781	95.36743164	
10										
11		105	1457	21315	326981	0	78.875	121.40625	1678.900391	
12										
13	md1	13.125								
14	md2	182.125								
15	md3	2664.375								
16	md4	40872.625								
17										
18	m1	0								
19	m2	9.859375								
20	m3	15.17578125								
21	m4	209.8625488								
22										
23		beta1	0.240300105							
24		gamma1	0.490204146	skewness						
25										
26		beta2	2.158918126							
27		gamma2	-0.84108187	kurtosis						
28										

RAW

Central

positively skewed

platykurtic

for grouped data.

$$\text{raw} = \frac{\Sigma f_i (x_i)^r}{\Sigma f_i}$$

$$\text{central} : \frac{\Sigma f_i (x_i - \bar{x})^r}{\Sigma f_i}$$

$$\text{constant} = \frac{\Sigma f_i (x_i - A)^r}{\Sigma f_i}$$

Q. AM = 5 $(m'_1 = 5)$

$$m_2 = 20 = \frac{\Sigma(x-5)^2}{n}$$

$$m_3 = 140 = \frac{\Sigma(x-5)^3}{n}$$

Find 3rd moment of distribut^n around 10.

$$\Rightarrow m_2 = m'_2 - m'^2_1$$

$$20 = m'_2 - 25$$

$$m'_2 = 45$$

$$m_3 = m'_3 - 3m'_1 m'_2 + 2m'^3_1$$

$$140 = m'_3 - 675 + 250$$

$$m'_3 = 565$$

$$m'_3(10) = \frac{\Sigma(x_i - 10)^3}{n}$$

$$= \frac{\Sigma x_i^3}{n} - 30\frac{\Sigma x_i^2}{n} + 300\frac{\Sigma x_i}{n} - 1000\frac{\Sigma n}{n}$$

$$= 565 - \underset{(1350)}{30(45)} + \underset{(1500)}{300(5)} - 1000$$

$$= -285$$

Q. first 2 moments of the distribution about the value 4 are -1.5 & 2.7

a) Find the moments around zero (1 & 2)

b) find mean & SD.

$$\Rightarrow \quad \frac{\Sigma (x_i - 4)^1}{n} = -1.5$$

$$\frac{\Sigma x_i}{n} - 4 = -1.5$$

$$\boxed{m'_1} = \frac{\Sigma x_i}{n} = \boxed{2.5} \ \boxed{\text{(mean)}}$$

$$\frac{\Sigma (x_i - 4)^2}{n} = 2.7$$

$$\frac{\Sigma x_i^2}{n} - 8\frac{\Sigma x_i}{n} + 16 = 2.7$$

$$\boxed{m'_2} = \frac{\Sigma x_i^2}{n} = 2.7 - 16 + 20 = \boxed{6.7}$$

$$\sigma = \frac{\Sigma (x_i - \bar{x})^2}{n} = \frac{\Sigma (x_i - 2.5)^2}{n}$$

$$\sigma = \frac{\Sigma x_i 2}{n} - 5\frac{\Sigma x_i}{n} + 6.25 = 6.7 - 12.5 + 6.25$$

$$\boxed{\text{SD}} = \sqrt{\sigma} = \sqrt{0.67} \qquad \boxed{= 0.45}$$

Q. $m'_1(1) = 2.6$

$m'_2(1) = 10.2$

$m'_3(1) = 43.4$

$m'_4(1) = 192.6$

Find AM & first 4 moments about 4

$$\Rightarrow m'_1 = \frac{\Sigma(x_i - 1)}{n}$$

$$2.6 = \frac{\Sigma x_i}{n} - 1$$

$$\frac{\Sigma x_i}{n} = 3.6 \quad \text{(mean)} \checkmark$$

$$m'_2 = \frac{\Sigma(x_i - 1)^2}{n}$$

$$10.2 = \frac{\Sigma x_i^2}{n} - 2\frac{\Sigma x_i}{n} + 1$$

$$10.2 + 2(3.6) - 1 = \frac{\Sigma x_i^2}{n}$$

$$16.4 = \frac{\Sigma x_i 2}{n}$$

$$m'_3 = \frac{\Sigma(x_i - 1)^3}{n}$$

$$43.4 = \frac{\Sigma x_i^3}{n} - 3\frac{\Sigma x_i^2}{n} + 3\frac{\Sigma x_i}{n} - 1$$

$$\frac{\Sigma x_i^3}{n} = 43.4 + 3(\underset{49.2}{16.4}) - 3(\underset{10.8}{3.6}) + 1$$

$$\frac{\Sigma x_i^3}{n} = 82.8$$

$$m'_4 = \frac{\Sigma(x_i - 1)^4}{n}$$

$$192.6 = \frac{\Sigma x_i^4}{n} - 4\frac{\Sigma x_i^3}{n} + 6\frac{\Sigma x_i^2}{n} - 4\frac{\Sigma x_i}{n} + 1$$

$$\frac{\Sigma x_i^4}{n} = 192.6 + 4(\underset{331.2}{82.8}) - 6(\underset{98.4}{16.4}) + 4(\underset{14.4}{3.6}) - 1$$

$$\frac{\Sigma x_i^4}{n} = 438.8$$

1 4 6 4 1
\+ − + − +

$$m'_1(4) = \frac{\Sigma(x_i - 4)}{n} = \frac{\Sigma x_i}{n} - 4$$

$$= 3.6 - 4$$

$$= -0.4 \checkmark$$

$$m'_2 = \frac{\Sigma(x_i - 4)^2}{n} = \frac{\Sigma x_i^2}{n} - 8\frac{\Sigma x_i}{n} + 16$$

$$= 16 \cdot 4 - 8(3 \cdot 6) + 16$$

$$= 3 \cdot 6 \checkmark$$

$$m'_3 = \frac{\Sigma(x_i - 4)^3}{n} = \frac{\Sigma x_i^3}{n} - 12\frac{\Sigma x_i^2}{n} + 48\frac{\Sigma x_i}{n} - 64$$

$$= 82 \cdot 8 - 12\underset{196.8}{(16 \cdot 4)} + 48\underset{172.8}{(3 \cdot 6)} - 64$$

$$= -5 \cdot 2 \checkmark$$

$$m'_4 = \frac{\Sigma(x_i - 4)^4}{n} = \frac{\Sigma x_i^4}{n} - 4\frac{\Sigma x_i^3 4}{n} + 6\frac{\Sigma x_i^2 4^2}{n} - 4\frac{\Sigma x_i 4^3}{n} + 4^4$$

$$\begin{matrix} 1 & 4 & 6 & 4 & 1 \\ + & - & + & - & + \end{matrix}$$

$$= 438 \cdot 8 - 16\underset{1324.8}{(82 \cdot 8)} + 96\underset{1574.4}{(16 \cdot 4)} - 256\underset{921.6}{(3 \cdot 6)} + 256$$

$$= 22 \cdot 8 \checkmark$$

Q. $x_i \rightarrow 0.\ 1.\ 2.\ 3.\ \ldots.\ k$

> $CF \rightarrow f_0.\ f_1.\ f_2.\ \ldots.\ f_k$

$$GT\ \ \bar{x} = \frac{\sum_{i=1}^{k} f_i}{n}$$

Q. $n = 10$

$\Sigma fx = -10$

$\Sigma fx^2 = 400$

$\Sigma fx^3 = -1000$

$\Sigma fx^4 = 5000$

Find first 4 moments about mean

$m'_1 = -1$

$m'_2 = 40$

$m'_3 = -100$

$m'_4 = 500$

$m_1 = 0$

$m_2 = m'_2 - m'^2_1 = 40 - 1 = 39$

$m_3 = m'_3 - 3m'_1 m'_2 + 2m'^3_1$

$= -100 - 3(-1)(40) + 2(-1)^3$

$= -100 + 120 - 2$

$= 18$

$m_4 = m'_4 - 4m'_1 m'_3 + 6m'_2 m'^2_1 - 3m'^4_1$

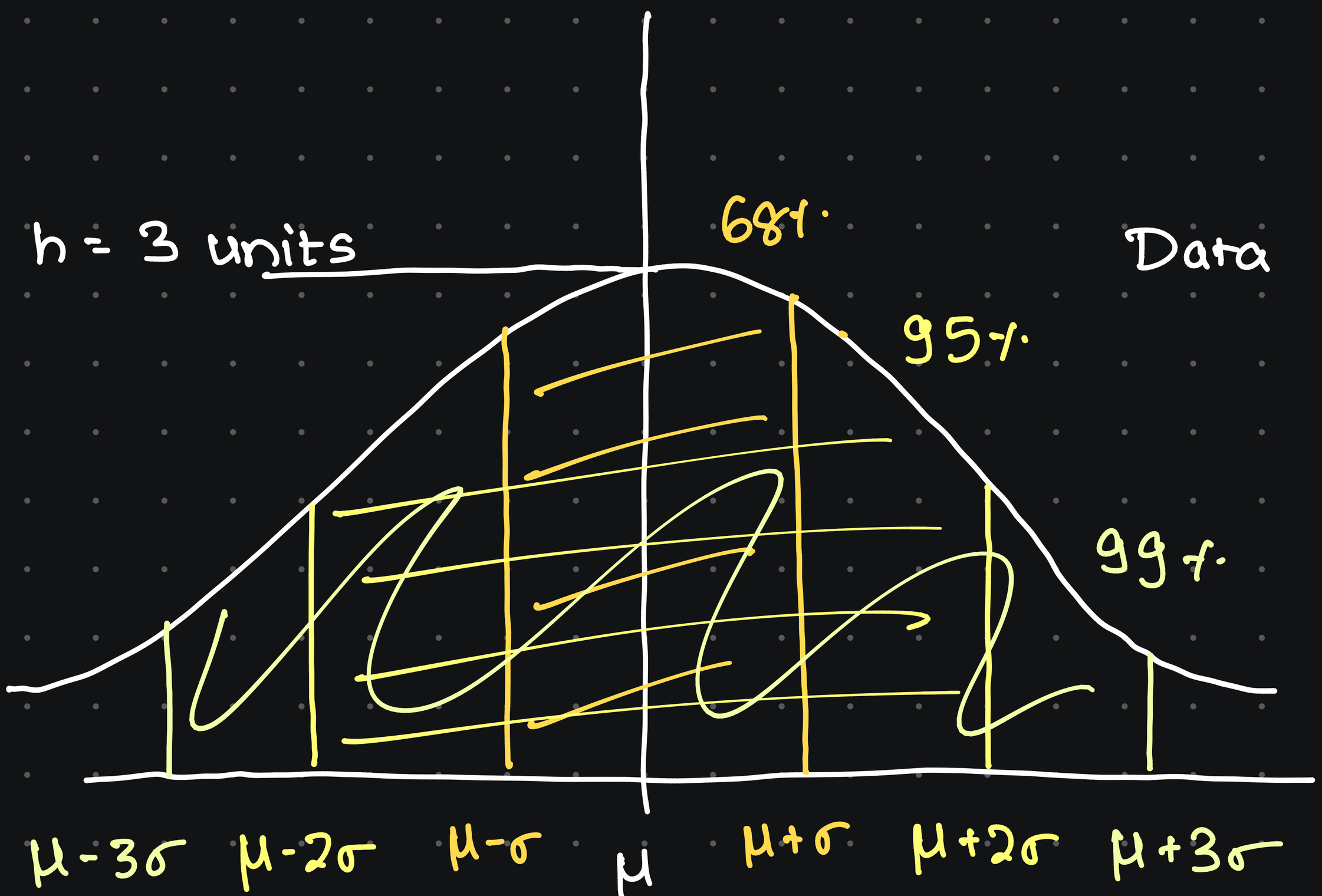
$= 500 - 4(-1)(-100) + 6(40)(1) - 3(-1)^4$

$= 500 - 400 + 240 - 3$

$= 337$

Normal Distribution

99 - 95 - 68



probability distributions

probability

Permutations

arrangement where order matters

$$nPr = \frac{n!}{(n-r)!}$$

$$= nCr \times r!$$

Combinations

selection (order does not matter)

$$nCr = \frac{n!}{(n-r)!\,r!}$$

$$= \frac{nPr}{r!}$$

$nC_0 = 1 \quad nC_1 = n$

$nC_n = 1$

$nC_r = nC_{n-r}$

$0! = 1$

Q. 1, 3, 7, 4, 5

now (4 digit nos without rep)?

$\Rightarrow$ $\underline{5}\ \underline{4}\ \underline{3}\ \underline{2} = \underline{\underline{120}}$

now (÷ by 5)

$\Rightarrow$ $\underline{4}\ \underline{3}\ \underline{2}\ \underline{\underline{1}}_{(5)} = \underline{\underline{24}}$

now (÷ by 2)

$\Rightarrow$ $\underline{4}\ \underline{3}\ \underline{2}\ \underline{\underline{1}}_{(4)} = \underline{\underline{24}}$

now (start with 7, end with 1)

$\Rightarrow$ $\underline{\underline{1}}_{(7)}\ \underline{3}\ \underline{2}\ \underline{\underline{1}}_{(1)} = \underline{\underline{6}}$

Q. 5 girls 4 boys 2 teachers photo

i) now (teachers not together)

ii) now (girls always together)

iii) now (boys —"—)

iv) now (girls separated)

⇒ total now = $(5+4+2)!$

$= 11!$

$= 39916800$

i) now (tr always 2gether)

$10! \times 2 = 7257600$ ↳ [11! − 10!2]

∴ now (not 2gether) = 32659200

ii) now = $7! 5! = 5040 \times 120 = 604800$

iii) now = $8! 4! = 40320 \times 24 = 967680$

iv) now = ↑B↑B↑B↑B↑T↑T↑

$= {}^{7}P_5 \times 6! = 2520 \times 720 = 1814400$

Q. 8 exam papers — 8 days
now (2 math papers together ?)
now (2 math separated ?)

⇒ total now = $8!$ = $\boxed{40320}$

let both math be together

∴ no. of papers = 6 + 1 = 7

i) ∴ now = $7!\,2!$ = 5040 × 2 = $\boxed{10080}$

ii) Separate

_ — _ — _ — _ — _ — _ — _

7 places for 2 math papers

now = $^7P_2$

∴ $\frac{7!}{5!} \times 6!$ = 42 × 720 = $\boxed{30240}$

OR

now = $8! - 7!\,2!$ = 40320 − 10080 = $\boxed{30240}$

Q. now ($\overset{1}{M}\overset{2}{A}\overset{3}{T}\overset{4}{H}\overset{5}{E}\overset{1}{M}\overset{2}{A}\overset{3}{T}\overset{9}{I}\overset{10}{C}\overset{11}{S}$)

M-2
A-2
T-2

now (begin with M)
now (begin with vowel)
now (end with consonant)
now (vowels together)
now (vowels separated)

i) now = $\frac{11!}{2!\,2!\,2!}$ = $\boxed{4989600}$

$(\underset{2}{M})\,(\underset{2}{A})\,(\underset{2}{T})$

repetitive : $\frac{n!}{m!\,p!\,r!\,\ldots.}$

ii) $1 \times \frac{10!}{2!\,2!}$ = $\boxed{907200}$

iii) E/I : $2 \times \frac{10!}{2!\,2!\,2!}$ = 907200

A : $1 \times \frac{10!}{2!\,2!}$ = 907200

$\therefore$ total now = $\boxed{1814400}$

iv) MTHCS

M : $1 \times \frac{10!}{2!\,2!}$ = 907200

$$T = 1 \times \frac{10!}{2!2!} = 907200$$

$$H/C/S = 3 \times \frac{10!}{2!2!2!} = 1360800$$

$$\therefore \text{now} = \boxed{3175200}$$

---

v) AAEI

$$\text{now}' = \frac{4!}{2!} = 12$$

$$\therefore \text{now} = 12 \times \frac{8!}{2!2!} = 12 \times \frac{40320}{4} = \boxed{120960}$$

vi) now =

_ _ _ _ _ _ _ _ _

$\frac{{}^8P_4}{2!}$ ⟶ vowels separated

$$\therefore \text{now} = \frac{{}^8P_4}{2!} \times \frac{7!}{2!2!}$$

8. now (CALCUTTA)
now(begin with vowel)
now (begin with vowel & end with consonant)
now(vowels together)
now(vowels separated)

⇒ CALCUTTA total = 8
1 1 2 2 1

C : 2
A : 2
T : 2

V	C
A	C
U	L
	T

i) now = $\frac{8!}{2!2!2!}$ = 5040

ii) begin with A: $1 \times \frac{7!}{2!2!}$

begin with U: $1 \times \frac{7!}{2!2!2!}$

∴ total now = $\frac{7!}{2!2!} + \frac{7!}{2!2!2!}$

iii) A C
A T
A L

U C
U T
U L
} same

selection

Combination : $nC_r = \frac{n!}{r!(n-r)!}$

$nC_1 = n = nC_{n-1}$

$nC_r = nC_{n-r}$

$nC_0 = 1$

Q. A box : 4R . 5W . 6B
3 selected Total = 15

now (exactly 2R)
now (atleast 2W)
now (atmost 1B)
now (all same color)
now (all diff color)

1. now (2R) = $4C_2 = \frac{4!}{2!2!} = 3! = 6$

now (1 out of 11) = $11C_1 = 11$

∴ total now = $6 \times 11 = \underline{66}$

2. i) 2W + 1

now (2 W) = $5C_2$
now (1 out of 10) = $10C_1$

∴ total now = $5C_2 \times 10C_1$
$= 10 \times 10$

ii) 3 W

now = $5C_3 = 10$

$\therefore$ ans: $(5C_2 \times 10C_1) + 5C_3$

$= 100 + 10$

$= \underline{110}$

3. i) 0 Black

$$\text{now} = 9C_3 = \frac{9!}{6!3!} = \frac{\overset{3}{\cancel{9}} \times 8 \times \overset{4}{7}}{\cancel{3} \times 2} = 84$$

ii) 1 B + 2

now = $6C_1 = 6$

$$\text{now} = 9C_2 = \frac{9!}{7!2!} = \frac{9 \times \overset{4}{\cancel{8}}}{\cancel{2}} = 36$$

$\therefore$ total now = $6 \times 36 = 216$

$\therefore$ ans : $84 + 216 = \underline{300}$

4. i) all B : $6C_3 = \frac{6!}{3!3!} = 20$

ii) all W = $5C_3 = 10$

iii) all R = $4C_3 = 4$

$\therefore$ total = $\underline{34}$

5. $^4C_1 \times {}^5C_1 \times {}^6C_1 = 4 \times 5 \times 6 = \underline{120}$

Q. 4 Cards
now (exact 2 R) ✓
now (exact 3 spades) –
now (atleast 3 clubs) –
now (no diamond) –
now (atleast 1 face) –
now (atmost 1 face) –
now (cards of diff. suits) –
now (cards of same suit) –
now (Q of ♡) –

1. 26 red
26 black

$$\text{now} = {}^{26}C_2 \times {}^{26}C_2 = 325 \times 325 = \underline{105625}$$

2. 13 spades
39

$^{13}C_3 \times {}^{39}C_1$

3. i) 3 clubs
$^{13}C_3 \times {}^{39}C_1$

ii) 4 clubs
$^{13}C_4$

$\therefore$ now $= {}^{13}C_3 \times {}^{39}C_1 + {}^{13}C_4$

4. no diamond

$39C_4$

5. i) 1 face

faces = 12

$12C_1 \times 40C_3$

ii) 2 face

$12C_2 \times 40C_2$

iii) 3 face

$12C_3 \times 40C_1$

iv) 4 face

$12C_4$

$\therefore$ now = $12C_4 + 12C_1\ 40C_3 + 12C_2 40C_2 + 12C_3 40C_1$

6. i) 0 face

$40C_4$

ii) 1 face

$12C_1 \times 40C_3$

$\therefore$ now = $40C_4 + 12C_1\ 40C_3$

7. diff suits — one from each

$$\text{now} = {}^{13}C_1 \cdot {}^{13}C_1 \cdot {}^{13}C_1 \cdot {}^{13}C_1$$

8. same suit — 4 from each, 4 suits

$$\text{now} = {}^{13}C_4 \times 4$$

9. Queen of hearts

$$\text{now} = {}^{1}C_1 \times {}^{51}C_2$$

1

1 1

$2C_0$ 1 $2C_1$ 2 $2C_2$ 1

$3C_0$ 1 $3C_1$ 3 $3C_2$ 3 $3C_3$ 1

1 $^{4}C_0$ 4 $^{4}C_1$ 6 $^{4}C_2$ 4 $^{4}C_3$ 1 $^{4}C_4$

1 $^{5}C_0$ 5 $^{5}C_1$ 10 $^{5}C_2$ 10 $^{5}C_3$ 5 $^{5}C_4$ 1 $^{5}C_5$

Combinatorics coefficients

Q. 3 sections in Q paper
each has 5 qs

ans 5 as
min 1 from each
now?

⇒

1 1 3 1 3 1 3 1 1	$5C_1$	$5C_1$	$5C_3$	250 × 3
	5	5	10	
1 2 2 2 1 2 2 2 1	$5C_1$	$5C_2$	$5C_2$	500 × 3
	5	10	10	

750 + 1500 = 2250

Q. 16 players
5 bowl
2 wk
9 bat
} team of 11

now(3 bowl)
now(atleast 2 bowl)
now (atleast 3 bowlers & 1 wk)
now(——— " ——— & atmost 1 wk)

1. $5C_3 \times 11C_8$

2. i) $5C_2 \times 11C_9$

ii) $5C_3 \times 11C_8$

iii) $5C_4 \times 11C_7$

iv) $5C_5 \times 11C_6$

} add

3. $5C_3 \times 2C_1 \times 9C_7 + 5C_4 \times 2C_1 \times 9C_6$

$+ 5C_5 \times 2C_1 \times 9C_5$

4. i) atl 3B. 0wk

$5C_3 \times 2C_0 \times 9C_8 + 5C_4 \times 2C_0 \times 9C_7 + 5C_5 \times 2C_0 \times 9C_6$

ii) atl 3B. 1 wk

$5C_3 \times 2C_1 \times 9C_7 + 5C_4 \times 2C_1 \times 9C_6 + 5C_5 \times 2C_1 \times 9C_5$

Experiment

Deterministic

↓

result is known &

confirming the result

Probabilistic

↓

result is unknown

↓

diff possibilities

random exp

sample

event

trials

mutually exclusive + exhaustive

outcome

population

set

unbiased

biased

gambling

conditional

Baye's thm

independent

variable

dependent

$$P(x) = \frac{\text{favorable outcomes}}{\text{total outcomes}}$$ } classified defn

for countable nos.

Q. 2 cards drawn from 52.
P (both red)
P (one ♡, one ♢)

⇒ 1. red = 26

$$\frac{26C_2}{52C_2}$$

2. one ♡, one ♢

$$\frac{13C_1 \cdot 13C_1}{52C_2}$$

Q. 10 articles. 4 defective.
3 chosen random
P(none defective)
└ 6 non-defect

⇒ $\frac{6C_3}{10C_3}$

Q. P(3 children in fam. diff bdays)

⇒ $\frac{365C_1 \cdot 364C_1 \cdot 363C_1}{365C_1 \cdot 365C_1 \cdot 365C_1}$

Sample Space

Finite

Put elements in sample space by one to one correspondence

infinite

countable infinite

⇓

one-to-one correspondence

eg. natural nos. & odd nos.

eg. coin tossed till head appears

$S = \{H, TH, TTH, \ldots\}$

uncountable infinite

defined by interval or union of intervals in real time

eg. $S = \{x : 0 < x < \omega\}$

infinite hotel rooms puzzle

count no. of leaves on a tree

count no. of stars in the universe

Write S.S for -

1] A die is rolled till 4 appears

$S = \{4, 14, 24, 34, 54, 64, 114, 124 \ldots\}$

2] 2 digit no. is formed from the digits 4,5,6,7 without rep.

$S = \{45, 46, 47, 54, 56, 57, 64, 65, 67, 74, 75, 76\}$

Sample Space

Discrete	Continous
definite values (not in interval)	$\{x \mid x \in \mathbb{Q} \text{ . st } 0 \leq x \leq 4\}$
$S = \{1, 2, \ldots, 10\}$	intervals
isolated values	uncountably infinite
countably finite/ infinite	

Events

5) Disjoint

$A \cap B = \phi$

1) Elementary event

s.s. with single element — singleton set

$S = \{a\}$

2) Impossible event

$S = \{\} / \phi \quad ; \quad P(e) = 0$

3) Sure event

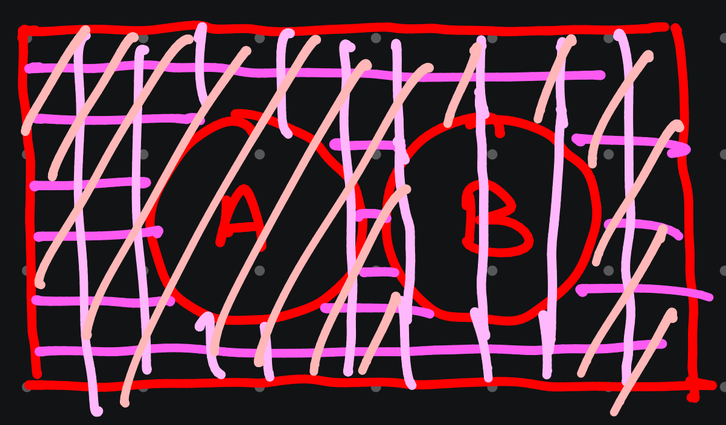
$P(e) = 1$

4) Complement (mutually exclusive + exhaustive)

$\Rightarrow A \cup A' = S$

D'morgan's Law

i) $\underline{(A \cup B)'} = \underline{A'} \cap \underline{B'}$



ii) $(A \cap B)' = A' \cup B'$

$\cup$: or	$\cap$: and

Golden Rules

- $A \cup B = B \cup A$
- $A \cup S = S$
- $A \cup A = A$
- $A \cup \phi = A$
- $A \cup A' = S$

} ∪

- $A \cap B = B \cap A$
- $A \cap S = A$
- $A \cap A = A$
- $A \cap \phi = \phi$
- $A \cap A' = \phi$

} ∩

only A : $A - A \cap B$ / $A \cap B'$

only B : $B - A \cap B$ / $A' \cap B$

Q. 3 coins tossed together

A : exactly 2 coins show H
B : atleast 2 coins — " —

Verify if A & B are mele?

⇒

T	T	T	
T	T	H	
T	H	T	
T	H	H	✓
H	H	H	✓
H	T	T	
H	T	H	✓
H	H	T	✓

$A : \frac{3}{8}$

$B : \frac{4}{8}$

not me

not e

$$n(A \cup B) = n(A) + n(B) - n(A \cap B)$$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$n(A \cup B \cup C) = n(A) + n(B) + n(C) - n(A \cap B) - n(B \cap C) - n(C \cap A) + n(A \cup B \cup C)$$

Q. 100 Customers

22% Suit ✓

30% Shirt ✓

28% tie ✓

11% Suit & shirt ✓

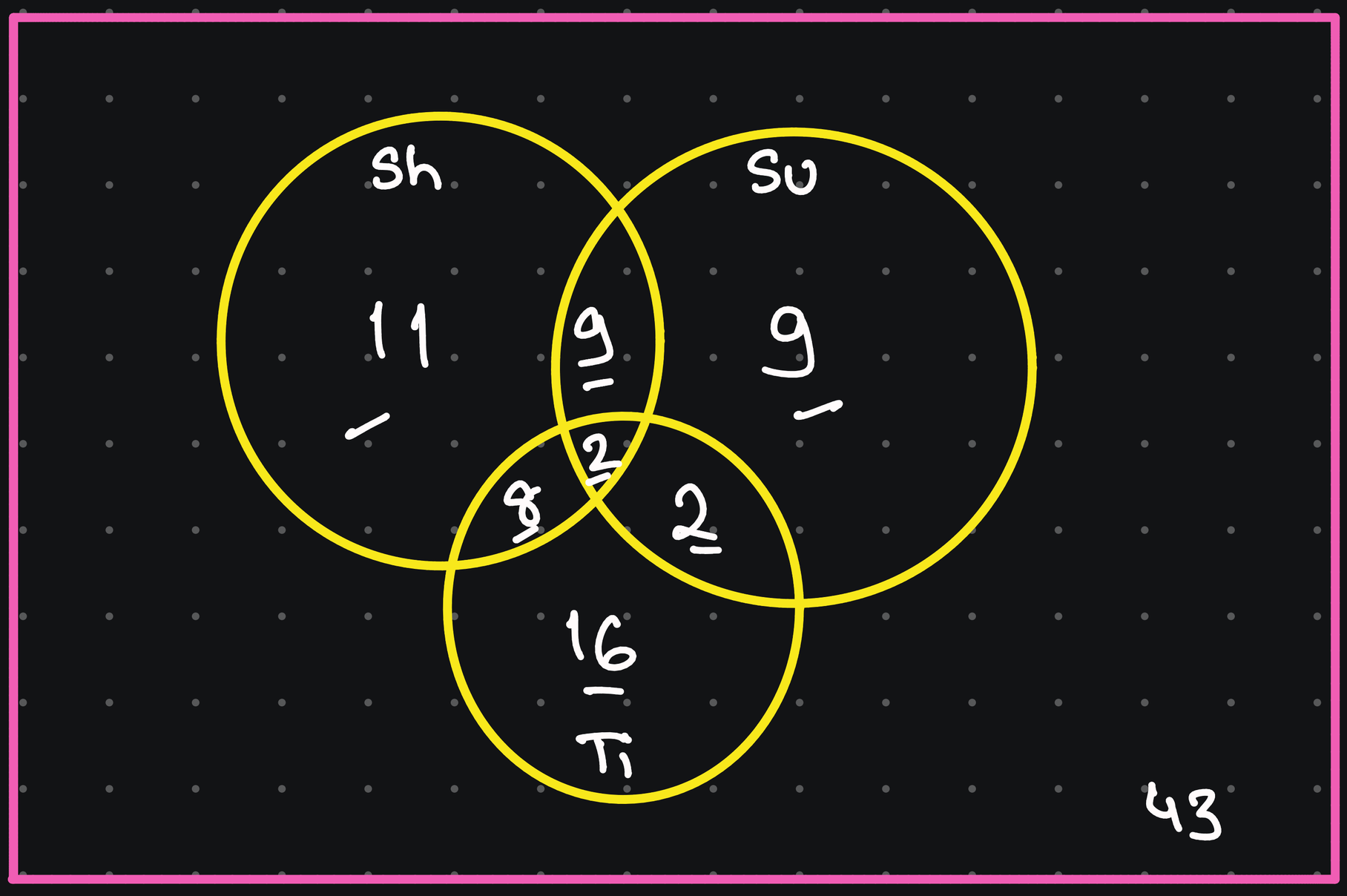
4% Suit & tie ✓

10% Shirt & tie ✓

2% all ✓

none? (43)

exactly one of 3? (36)



Q. 4 cards from 52

$$P(\text{all spades}) \Rightarrow \frac{13C_4}{52C_4}$$

$$P(\text{all diff suits}) \Rightarrow \frac{13C_1 \cdot 13C_1 \cdot 13C_1 \cdot 13C_1}{52C_4}$$

$$\left(\frac{13C_1{}^4}{52C_4}\right)$$

(*) $$P(\text{K.J.A, random}) \Rightarrow \frac{4C_1 \cdot 4C_1 \cdot 4C_1 \cdot 49C_1}{52C_4}$$

Q. 2 die thrown

P(sum on uppermost face > 8)

total = 36

3.6 ; 4.5 ; 4.6 ; 5.4 ; 5.5 ; 5.6 ; 6.3 ; 6.4 ; 6.5 ; 6.6

$$P = \frac{10}{36} = \frac{5}{18}$$

P(sum is neither 7 nor 11)

	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

$$P = \frac{28}{36} = \frac{7}{9}$$

Q. Elevator : 5 passengers
stops @ 9 floors

P(no 2 passengers get down at same floor)

assume: every floor has equal prob. of passengers.

$\Rightarrow$ now $= 9^5$ $\quad \underline{9}\ \underline{9}\ \underline{9}\ \underline{9}\ \underline{9}$

favorable $= {}^9P_5$

$\therefore P = \dfrac{{}^9P_5}{9^5}$

Q. 2 R
3 Bu
5 Bc
3 drawn at random

$P(\text{all 3 diff colors}) = \dfrac{{}^2C_1 \cdot {}^3C_1 \cdot {}^5C_1}{{}^{10}C_3}$

P(exactly 2 of same color)

i) $2R : \dfrac{{}^2C_2 \cdot {}^8C_1}{{}^{10}C_3}$ +

ii) $2Bu = \dfrac{{}^3C_2 \cdot {}^7C_1}{{}^{10}C_3}$ +

iii) $2Bc = \dfrac{{}^5C_2 \cdot {}^5C_1}{{}^{10}C_3}$

9. COMPUTER

i. P(vowels occupy even places)
ii. P(vowels occupy odd places)
iii. P(vowels together)
iv. P(vowels separated)

⇒ letters = 8
vowels = O. U. E = 3

total possibilities (d) = 8!

i) vowels @ even places
vowels : $^4P_3$
consonants : 5!

$$\therefore P_i = \frac{^4P_3 \cdot 5!}{8!} = \frac{\cancel{24}^{4!} \cdot 5!}{\cancel{8}^{2} \cdot 7 \cdot \cancel{6} \cdot \cancel{5!}} = \boxed{\frac{1}{14}}$$

ii) as no. of even places = no. of odd places.

$$P_{ii} = \frac{^4P_3 \cdot 5!}{8!} = \boxed{\frac{1}{14}}$$

iii) OUE together
now = 3!

remaining (5 cons + OUE) = 6!

$$\therefore P_{iii} = \frac{3!\,6!}{8!} = \frac{\cancel{6}^{3} \cdot 6!}{\cancel{8} \cdot 7 \cdot 6!} = \boxed{\frac{3}{28}}$$

iv) $\text{L C L C L C L C L C L}$

$\therefore$ now $= {}^{6}P_3 \cdot 5!$

$\therefore P_{iv} = \dfrac{{}^{6}P_3 \cdot 5!}{8!} = \dfrac{6!\,5!}{3!\,8!} = \dfrac{\cancel{6!}\cdot \overset{20}{\cancel{120}}^{5}}{\underset{2}{\cancel{6}\cdot\cancel{8}\cdot 7\cdot\cancel{8!}}}$

ans. $= \boxed{\dfrac{5}{14}}$

Classical defn of probability

$P(e) = \dfrac{m}{n}$ —— favorable
—— total

assumpt^n : every event in 'm' is equally likely

drawback ⟶ random expt. doesn't always result in equally likely outcomes

death rate, child birth ratio etc. —— Vital Stats

useful for insurance policies

not suitable for infinitely many outcomes
eg. coin tossed till head appears

relative approach of probability

$$P(A) = \lim_{N \to \infty} \frac{m}{n}$$

if limit exists

> difficult to maintain ideal & homogenous conditions

(drawback)

> $\frac{m}{n}$ may not attain a unique value irrespective of large 'n'

* $P\{\omega_i\} \geq 0$

$$\sum_{i=1}^{s} P(\omega_i) = 1$$

Q. $S = \{\omega_1, \omega_2, \omega_3\}$

$P(\omega_1) = \frac{5}{7}$ $P(\omega_2) = \frac{3}{14}$ $P(\omega_3) = \frac{1}{7}$

prob. model?

$\Rightarrow \frac{5}{7} + \frac{3}{14} + \frac{1}{7} = \frac{15}{14} \neq 1$

$\therefore$ not a prob. model.

Q. $\mathbb{Z}$ chosen from 1-100. both inclusive.

Prob(selecting perfect square) if

i. all integers equally likely

ii. $\mathbb{Z}$ b/lw 1-50 twice as likely to occur than an $\mathbb{Z}$ bllw 51-100.

$\Rightarrow$ i) $P(x) = \frac{10}{100} = \boxed{\frac{1}{10}}$

ii) 1-50: 7 squares $P(A) \rightarrow 2P$
51-100: 3 squares $P(B) \rightarrow P$ } $3P$

$\frac{7}{50}\cdot\frac{2}{3} + \frac{3}{50}\cdot\frac{1}{3} = \frac{14}{150} + \frac{3}{150} = \boxed{\frac{17}{150}}$

$P(A) = 2P$ 50

$P(B) = P$ 50

$50 \cdot 2P + 50 \cdot P = 1$

$P = \frac{1}{150}$

$P(x) = 7 \cdot 2P + 3 \cdot P$

$= 14P + 3P$

$= 17P$

ans. $\boxed{= \frac{17}{150}}$

Q. 52 cards

$P(\text{red}) = 3\,P(\text{black})$ $\therefore R = \frac{3}{4}$

$B = \frac{1}{4}$

i. P(red face)

$\frac{6^3}{26} \cdot \frac{3}{4_2} = \frac{9}{52}$

26 26 26 26

R B

6 6 6

Q. win I : P
4 : 1
out of 20, I wins?

$\Rightarrow P(I) = \frac{4}{5}$

$\therefore$ win $= \frac{4}{5} \times 20 = \boxed{16}$

$\therefore P(I \text{ wins}) = \frac{16}{20}$

Axiomatic Approach

↳ accepted without proof

1. P(A) is a real number.
2. $P(A) \geq 0$ for all sample spaces
3. $P(S) = 1$
4. $P(A \cup B) = P(A) + P(B)$ if mutually exclusive events

Axiom 1 : $P(A) \geq 0$ for any event A on S.
Axiom 2 : $P(S) = 1$
Axiom 3 : $P(A \cup B) = P(A) + P(B)$ for every pair of mutually exclusive events A & B defined on S.

Theorem 1: For any event A on S
$0 \leq P(A) \leq 1$

By axiom 1, $P(A) \geq 0$
$A \cup A' = S$
$A \cap A' = \phi$
A and A' are disjoint on S
$P(A \cup A') = P(A) + P(A')$
$1 = P(A) + P(A')$
But $P(A') \geq 0$
$1 - P(A) \geq 0$
$P(A) \leq 1$

Theorem 2:
To prove that, $P(\phi) = 0$
Let A be an event on S
$A \cup \phi = A$, $A \cap \phi = \phi$
A & ϕ are mutually exclusive
$\therefore P(A \cup \phi) = P(A) + P(\phi)$, Using Axiom 3
$\therefore P(A) = P(A) + P(\phi)$
$\therefore P(\phi) = 0$

t A on S.

every pair of
e events
on S.

s

iom 3

v.v. imp.

Theorem 3 : A & B are defined on S such that $A \subset B$, then $P(A) \leq P(B)$

$B = A \cup (A' \cap B)$

$P(B) = P[A \cup (A' \cap B)]$

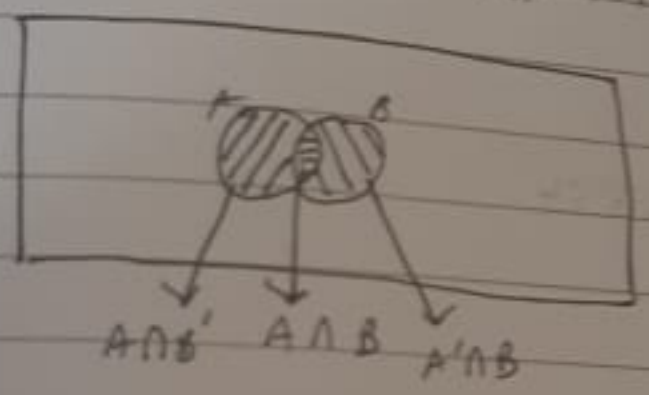
$= P(A) + P(A' \cap B)$

Now $P(A' \cap B) \geq 0$

$\therefore P(B) \geq P(A)$

Theorem 4 : Addition Theorem of Probability / Union

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$



$P(A \cup B) = P(A \cap B') + P(A \cap B) + P(A' \cap B)$

But, $P(A) = P(A \cap B') + P(A \cap B)$

$P(B) = P(A' \cap B) + P(A \cap B)$

$\therefore P(A \cup B) = P(A) - P(A \cap B) + P(B) - P(A \cap B) + P(A \cap B)$

$= P(A) + P(B) - P(A \cap B)$

6. $P(A) = 0.5$ $\quad P(A \cup B) = ?$

$P(B) = 0.7$ $\quad P(A' \cap B') = ?$

$P(A \cap B) = 0.3$

Sol^n $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

$= 0.9$

$P(A') = 0.5$

$P(B') = 0.3$

De Morgan's law

$P(A' \cap B') = 1 - P(A \cup B) = P(A \cup B)'$

$= 1 - 0.9$

$= 0.1$

8 $P(A) = 5P(A)'$

$= 5 - 5P(A)$

$6P(A) = 5$

$P(A) = \frac{5}{6} = 0.834$

07/11/2023

* Let A & B, two events define on S, such that $P(A) = \frac{3}{4}$ and $P(B) = \frac{5}{8}$.

P.T ① $\frac{3}{8} \leq P(A \cap B) \leq \frac{5}{8}$

② $P(A \cup B) \geq \frac{3}{4}$

③ $\frac{1}{8} \leq P(A \cap B') \leq \frac{3}{8}$

Ans:

~~P(A~~ $(A \cap B) \subset A$, $A \cap B \subset B$.

$P(A \cap B) \leq P(A)$, $P(A \cap B) \leq P(B)$

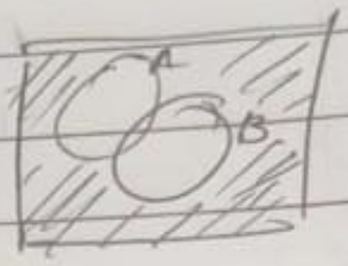
$\therefore P(A \cap B) \leq \min\left\{\frac{3}{4}, \frac{5}{8}\right\}$.

① $\therefore P(A \cap B) \leq \frac{5}{8}$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$P(A \cap B) = P(A) + P(B) - P(A \cup B)$$

$$P(A \cap B) \geq P(A) + P(B) - 1$$



$$P(A \cap B) \geq \frac{3}{4} + \frac{5}{8} - 1$$

$$P(A \cap B) \geq \frac{3}{8} \quad -(2)$$

from (1) & (2):

$$\boxed{\frac{3}{8} \leq P(A \cap B) \leq \frac{5}{8}}$$

P-1: $P(A \cup B) \geq \frac{3}{4}$

$A \subseteq A \cup B$, $B \subseteq A \cup B$

$P(A) \leq P(A \cup B)$, $P(B) \leq P(A \cup B)$

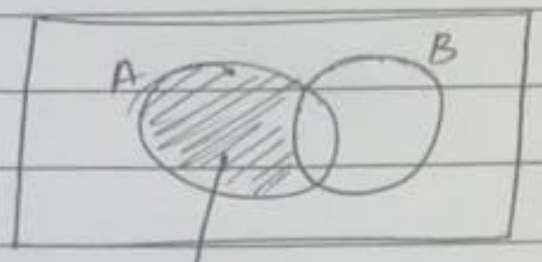
$$P(A \cup B) \geq \max\{P(A), P(B)\}$$

$$P(A \cup B) \geq \max\left\{\frac{3}{4}, \frac{5}{8}\right\}$$

$$P(A \cup B) \geq \frac{3}{4}$$ // Hence proved.

PAGE No
DATE : / 202

(3) P.T: $\frac{1}{8} \leq P(A \cap B') \leq \frac{3}{8}$



$(A \cap B') \subset A$. $\quad P(A \cap B') \leq P(A)$

$(A \cap B') \subset B'$ $\quad P(A \cap B') \leq P(B')$

$P(A \cap B') \leq \min\{P(A), P(B')\}$

$P(A \cap B') \leq \min\{\frac{3}{4}, \frac{3}{8}\}$

$$\boxed{P(A \cap B') \leq \frac{3}{8}.} \quad -(4)$$

$P(A \cup B') = P(A) + P(B') - P(A \cap B')$

$P(A \cap B') \geq P(A) + P(B') - 1$

$P(A \cap B') \geq \frac{3}{4} + \frac{3}{8} - 1$

$$\boxed{P(A \cup B') \geq \frac{1}{8}} \quad -(5)$$

from (4) & (5),

$$\boxed{\frac{1}{8} \leq P(A \cap B') \leq \frac{3}{8}}$$

Hence Proved

PAGE No.
DATE: / 202

* What is the probability that a leap year selected at random will contain either 53 thursday or 53 fridays.

(Union) Ans: Mon, Tue
Tue, Wed
Wed, Thus
Thurs, Fri
Fri, Sat
Sat, Sun
Sun, Mon.

$\frac{52}{7}$
$\frac{7}{364}$

Let A be the event that leap year consists of 53 thursdays.

Let B be the event that leap year consists of 53 fridays.

$P(A) = \frac{2}{7}$

$P(B) = \frac{2}{7}$.

$P(A \cap B) = \frac{1}{7}$.

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$

$= \frac{2}{7} + \frac{2}{7} - \frac{1}{7}$

$= \frac{3}{7}$.

Singaporean Chaos.

PAGE No.
DATE: / 202

numbered.

* A bag contains 30 tickets from 1-30. 1 ticket is selected at random. Find the probability that number is either odd or the square of an integer.

$n(S) = 30$.

Let A be the event that number selected is odd

Let B be the event that number selected is square of integer.

$A = \{1, 3, 5 \cdots 29\}$
$n(A) = 15$

$B = \{1, 4, 9, 16, 25\}$
$n(B) = 5$

$P(A) = \frac{15}{30} = \frac{1}{2}$. $P(B) = \frac{5}{30} = \frac{1}{6}$.

$P(A \cap B) = \frac{3}{30} = \frac{1}{10}$.

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$

$= \frac{15}{30} + \frac{5}{30} - \frac{3}{30} = \frac{17}{30}$

$= \frac{17}{30} = \frac{85}{30} = \frac{3}{2}$

* Conditional Probability:

↳ Sampling without Replacement

$$P(A \cap B) = P(A) \cdot P(B|A)$$

read as prob of B given A

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

some outcome occurs given some other outcome has already occurred

① ↳ prob(B) depends on prob(A) happening

②

* Why are some mutual funds managers more successful than others.

	B_1: M.F outperforms market	B_2: M.F. does not O.P. market	
A_1: Top 20 MBA Programs	0.11	0.29	0.4
A_2: Not top 20.	0.06	0.54	0.6
	0.17	0.83.	1

marginal Prob

$P(A_1 \cap B_1) = 0.11$

$P(A_1 \cap B_2) = 0.29$

Marginal Prob. / Unconditional probability

$P(A_2 \cap B_1) = 0.06$

$P(A_2 \cap B_2) = 0.54$

Os: We select 1 mutual fund at random and discover that it did not outperform the market.

What is the probability that a graduate of top 20 MBA program manages it.

Ans:

$P(A_1|B_2) = ?$ = ~~0.29~~

$= \frac{P(A_1 \cap B_2)}{P(B_2)}$

$= \frac{0.29}{0.83} = 0.34.$

=> Are A_1 & B_1 independent?

$P(A_1|B_1) = \frac{P(A_1 \cap B_1)}{P(B_1)} = \frac{0.11}{0.17} = 0.647.$

$P(A_1) = 0.4.$

If they are independant, $P(A_1) = P(A_1|B_1)$

$$P(A|B) = \frac{P(A)}{P(B)} \cdot P(B|A)$$

8. $P(A \cup B) = \frac{7}{8}$

$P(A \cap B) = \frac{1}{4}$

$P(B') = \frac{1}{4}$

Find i) $P(B)$ iii) $P(A' \cap B')$ v) $P(A' \cap B)$

ii) $P(A)$ iv) $P(A' \cup B')$

$\Rightarrow$ i) $P(B) = 1 - P(B') = 1 - \frac{1}{4} = \boxed{\frac{3}{4}}$

ii) $P(A) + P(B) - P(A \cap B) = P(A \cup B)$

$\therefore P(A) = P(A \cup B) + P(A \cap B) - P(B)$

$= \frac{7}{8} + \frac{1}{4} - \frac{3}{4}$

$= \frac{7}{8} - \frac{2}{4} \frac{4}{8}$

$= \boxed{\frac{3}{8}}$

iii) $P(A' \cap B') = P(A \cup B)' = 1 - P(A \cup B)$

$$= 1 - \frac{7}{8}$$

$$= \boxed{\frac{1}{8}}$$

iv) $P(A' \cup B') = P(A \cap B)' = 1 - P(A \cap B)$

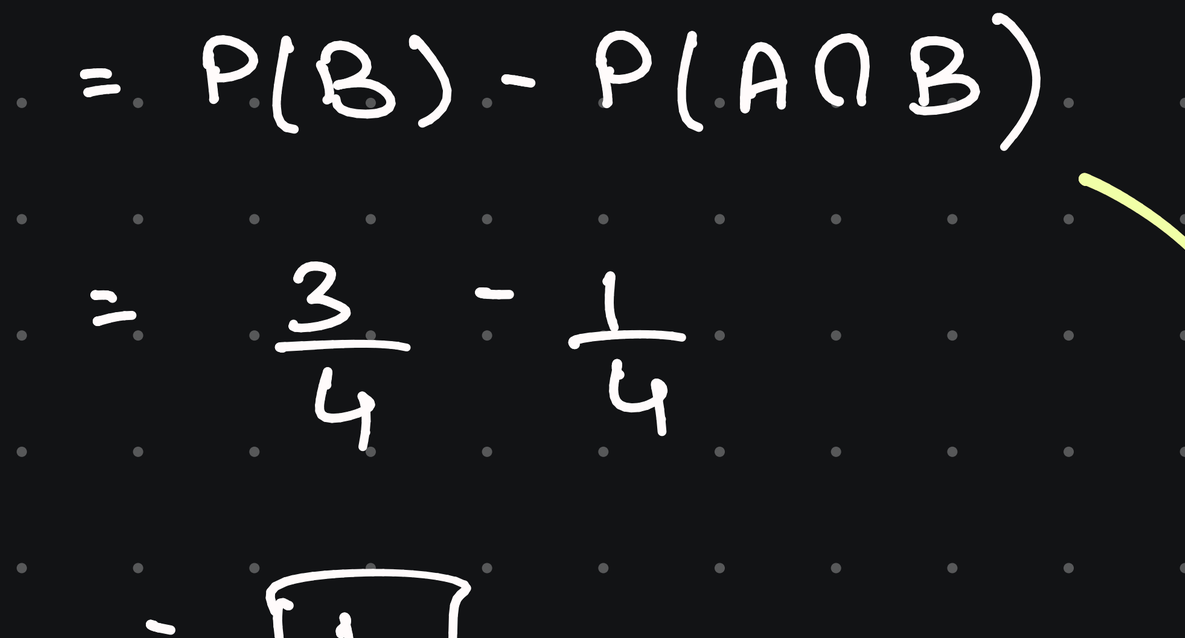
$$= 1 - \frac{1}{4}$$

$$= \boxed{\frac{3}{4}}$$

v) $P(A' \cap B) = P(B) - P(A \cap B)$

$$= \frac{3}{4} - \frac{1}{4}$$

$$= \boxed{\frac{1}{2}}$$



8. 30 stars
24 maths
3 selected at random

i) P(atleast 1 math chosen)

ii) P(none of math chosen)

$\Rightarrow$ now(select 3) = $(30+24)C_3 = 54C_3$

now(no math) = $30C_3$ (all stats)

$\therefore$ i) P(atleast 1) = 1 - P(no)

$= 1 - \frac{30C_3}{54C_3}$

$= 1 - 0.164$

$= \boxed{0.836}$

ii) $P(no) = \frac{30C_3}{54C_3} = \boxed{0.164}$

Conditional Probability

$$P(A \cap B) = P(A) \cdot P(B|A)$$

$$= P(B) \cdot P(A|B)$$

#For independent events,
$P(B|A) = P(B)$ &
$P(A|B) = P(A)$

$$\therefore P(B|A) = \frac{P(A \cap B)}{P(A)}$$

and
$P(A \cap B) = P(A) \cdot P(B)$

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

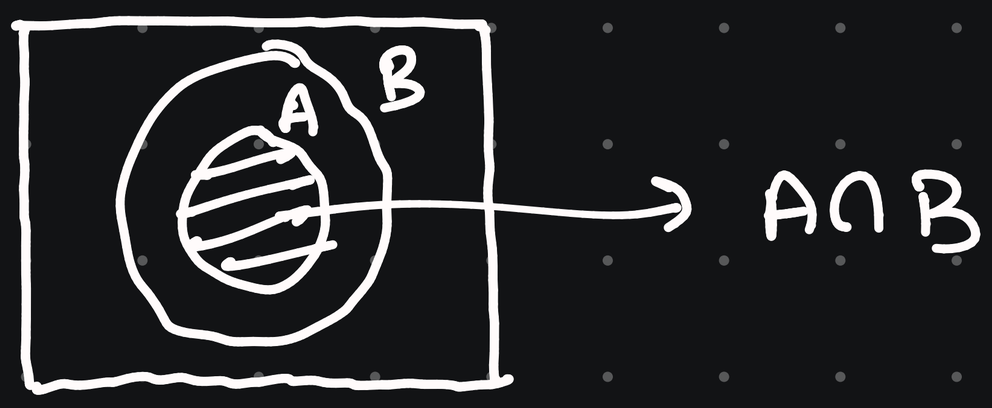
i) $P(A'|A) = 0 = \frac{P(A' \cap A)}{P(A)} = \frac{0}{P(A)} = 0$

ii) If A & B are mutually exclusive,

then $P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{0}{P(B)} = 0$

iii) If $A \subset B$ then

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{P(A)}{P(A)} = 1$$



Q. A → 1 out of 3 problems

B → 1 4

C → 20% problems

i) P(all of them solve the problem)

ii) P(only A solves)

iii) P (only 1 solves)

iv) P(only A & B solve)

v) P(A & B solve)

vi) P (neither solve)

vii) P(problem solved)

⇒ $P(A) = \frac{1}{3}$ $P(A') = \frac{2}{3}$

$P(B) = \frac{1}{4}$ $P(B') = \frac{3}{4}$

$P(C) = \frac{20}{100} = \frac{1}{5}$ $P(C') = \frac{4}{5}$

i) $P(\text{all solve}) = P(A \cap B \cap C)$

$= P(A) \cdot P(B) \cdot P(C)$

$= \frac{1}{3} \cdot \frac{1}{4} \cdot \frac{1}{5}$

$= \boxed{\frac{1}{60}}$

ii) $P(\text{only } A) = P(A \cap B' \cap C')$

$= P(A) \cdot P(B') \cdot P(C')$

$= \frac{1}{3} \cdot \frac{3}{4} \cdot \frac{4}{5}$

$= \boxed{\frac{1}{5}}$

iii) $P(\text{only 1}) = P(\text{only } A) \text{ or } P(\text{only } B) \text{ or } P(\text{only } C)$

$= P(A \cap B' \cap C') + P(B \cap A' \cap C') + P(C \cap B' \cap C')$

$= \frac{1}{3} \cdot \frac{3}{4} \cdot \frac{4}{5} + \frac{1}{4} \cdot \frac{2}{3} \cdot \frac{4}{5} + \frac{1}{5} \cdot \frac{2}{3} \cdot \frac{3}{4}$

$= \frac{1}{5} + \frac{2}{15} + \frac{1}{10} = \frac{6+4+3}{30} = \boxed{\frac{13}{30}}$

iv) $P(\text{only A \& B}) = P(A \cap B \cap C')$

$= P(A) \cdot P(B) \cdot P(C')$

$= \frac{1}{3} \cdot \frac{1}{4} \cdot \frac{4}{5}$

$= \boxed{\frac{1}{15}}$

v) $P(A \text{ \& } B \text{ solve}) = P(A \cap B \cap C) \text{ or } P(A \cap B \cap C')$

$= \frac{1}{60} + \frac{1}{15}$

$= \frac{1+4}{60}$

$= \boxed{\frac{1}{12}}$

vi) $P(\text{neither}) = P(A' \cap B' \cap C')$

$= \frac{2}{3} \cdot \frac{3}{4} \cdot \frac{4}{5}$

$= \boxed{\frac{2}{5}}$

vii) $P(\text{solved}) = 1 - P(\text{neither}) = \boxed{\frac{3}{5}}$

Concept of Odds

1. Odds in favor of A is m:n

$$P(A) = \frac{\text{fav.}}{\text{total}} = \frac{m}{m+n}$$

2. Odds against A is m:n

$$P(A) = \frac{\text{fav.}}{\text{total}} = \frac{n}{m+n}$$

Q. odds in favor of A = 3:4

$$\therefore P(A) = \frac{3}{7}$$

Q. odds against B = 5:3

$$\therefore P(B) = \frac{3}{8}$$

Q. odds in favor of person A speaking truth is 2:3

odds against B speaking truth = 5:1

$\therefore P(A) = \frac{2}{5}$ (truth) $P(B) = \frac{1}{6}$

i) P(tell same thing?)

ii) P(contradict?)

$\Rightarrow P(A') = \frac{3}{5}$ (false) $P(B') = \frac{5}{6}$

i) P(tell same) = P(both true) or P(both False)

$= P(A \cap B)$ or $P(A' \cap B')$

$= \frac{\cancel{2}^{1}}{5} \cdot \frac{1}{\cancel{3}} + \frac{\cancel{3}^{1}}{\cancel{5}} \cdot \frac{\cancel{5}}{\cancel{6}_{2}}$

$= \frac{1}{15} + \frac{1}{2}$

$= \boxed{\frac{17}{30}}$

ii) P(contradict) = 1 - P(same) = $\boxed{\frac{13}{30}}$

8. odds in favor of H = 1:4
odds against W = 3:4

i) P(atleast one)
ii) P(none)

$\Rightarrow P(H) = \frac{1}{5}$ $\quad P(W) = \frac{4}{7}$

$P(H') = \frac{4}{5}$ $\quad P(W') = \frac{3}{7}$

i) $P(\text{atleast one}) = P(A' \cap B) + P(A \cap B') + P(A \cap B)$

$= \frac{4}{5} \cdot \frac{4}{7} + \frac{1}{5} \cdot \frac{3}{7} + \frac{1}{5} \cdot \frac{4}{7}$

$= \frac{16}{35} + \frac{3}{35} + \frac{4}{35}$

$= \boxed{\frac{23}{35}}$

OR $P(\text{atl. one}) = 1 - P(\text{none}) = \boxed{\frac{23}{35}}$

ii) $P(\text{none}) = P(A' \cap B')$

$= \frac{4}{5} \cdot \frac{3}{7}$

$= \boxed{\frac{12}{35}}$

Types of Selections

1. Simultaneous
2. One-by-one ——— 2 diff events

1

With replacement ⇊ independent

Without replacement ⇊ conditional

Q. 4 math 6 eng
draw 2. one by one
i) with replacement
ii) without replacement

i. P(both math)
ii. P(first math second eng)
iii. P(one eng one math)
iv. P(both eng)

⇒ a) With replacement b) Without repl.

And

i. $\frac{4C_1}{10C_1} \cdot \frac{4C_1}{10C_1} = \frac{16}{100}$

i. $\frac{4C_1}{10C_1} \cdot \frac{3C_1}{10C_1} = \frac{12}{90}$

ii. $\frac{4C_1}{10C_1} \cdot \frac{6C_1}{10C_1}$

ii. $\frac{4C_1}{10C_1} \cdot \frac{6C_1}{9C_1}$

iii. $\frac{4C_1}{10C_1} \cdot \frac{6C_1}{10C_1} + \frac{6C_1}{10C_1} \cdot \frac{4C_1}{10C_1}$

iv. $\frac{6C_1}{10C_1} \cdot \frac{6C_1}{10C_1}$

iii. $\frac{4C_1}{10C_1} \cdot \frac{6C_1}{9C_1} + \frac{6C_1}{10C_1} \cdot \frac{4C_1}{9C_1}$

iv. $\frac{6C_1}{10C_1} \cdot \frac{6C_1}{9C_1}$

P(E∩m) or P(m∩E)

Q. 4 cards drawn 1 by 1.

a) without b) with repl.

P(first 2 Aces, next 2 queens)

⇒ a) $\frac{4C_1}{52C_1} \cdot \frac{3C_1}{51C_1} \cdot \frac{4C_1}{50C_1} \cdot \frac{3C_1}{49C_1}$

b) $\frac{4C_1}{52C_1} \cdot \frac{4C_1}{52C_1} \cdot \frac{4C_1}{52C_1} \cdot \frac{4C_1}{52C_1}$

2 Stage Problems
(draw probability tree)

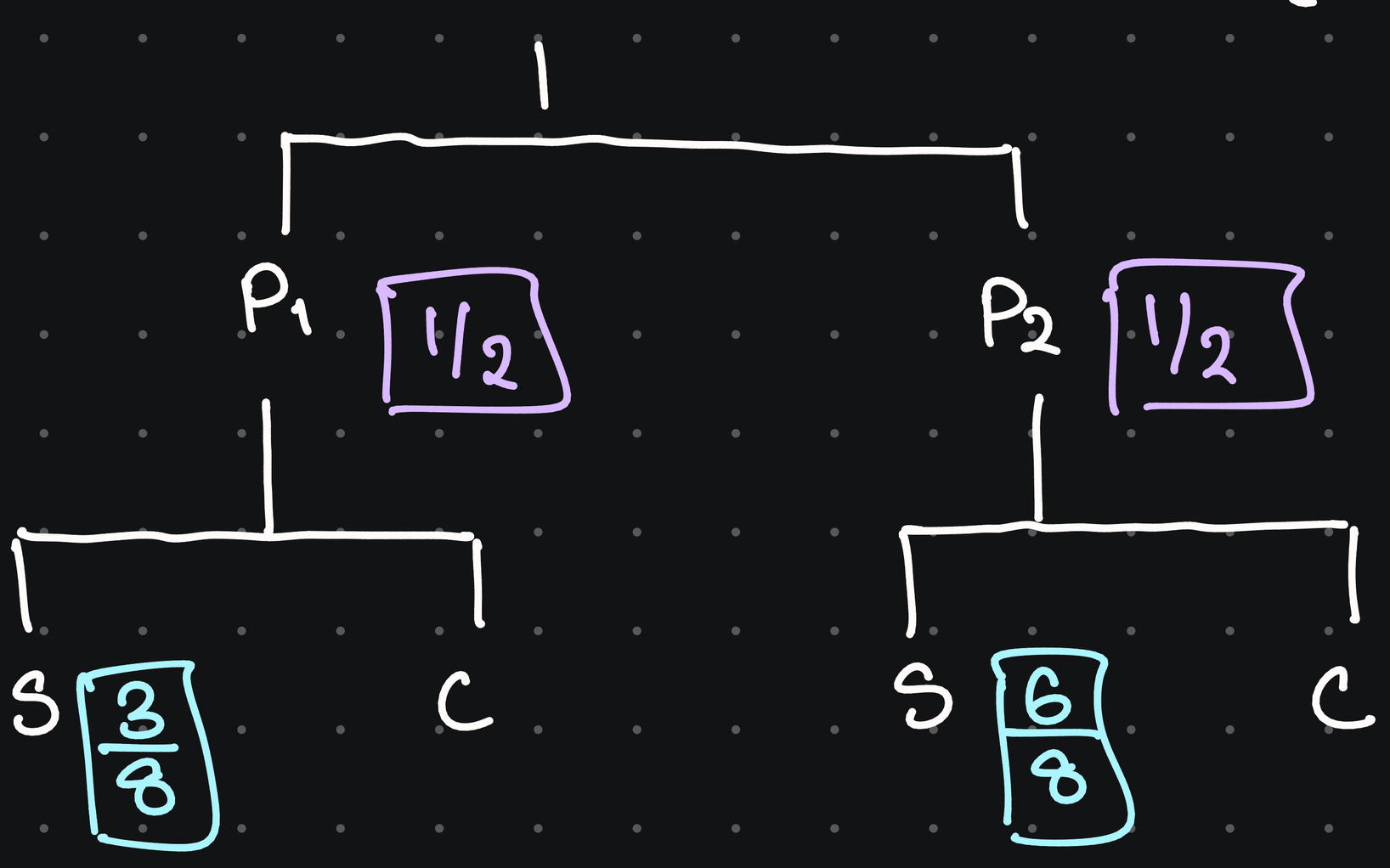
Q. Purse 1: 3 silver
5 copper

Purse 2: 6 silver
2 copper

Purse drawn at random & then coin is selected.

i) P(coin is silver)

ii) P(coin copper)
1 - P(silver)
$= \frac{7}{16}$

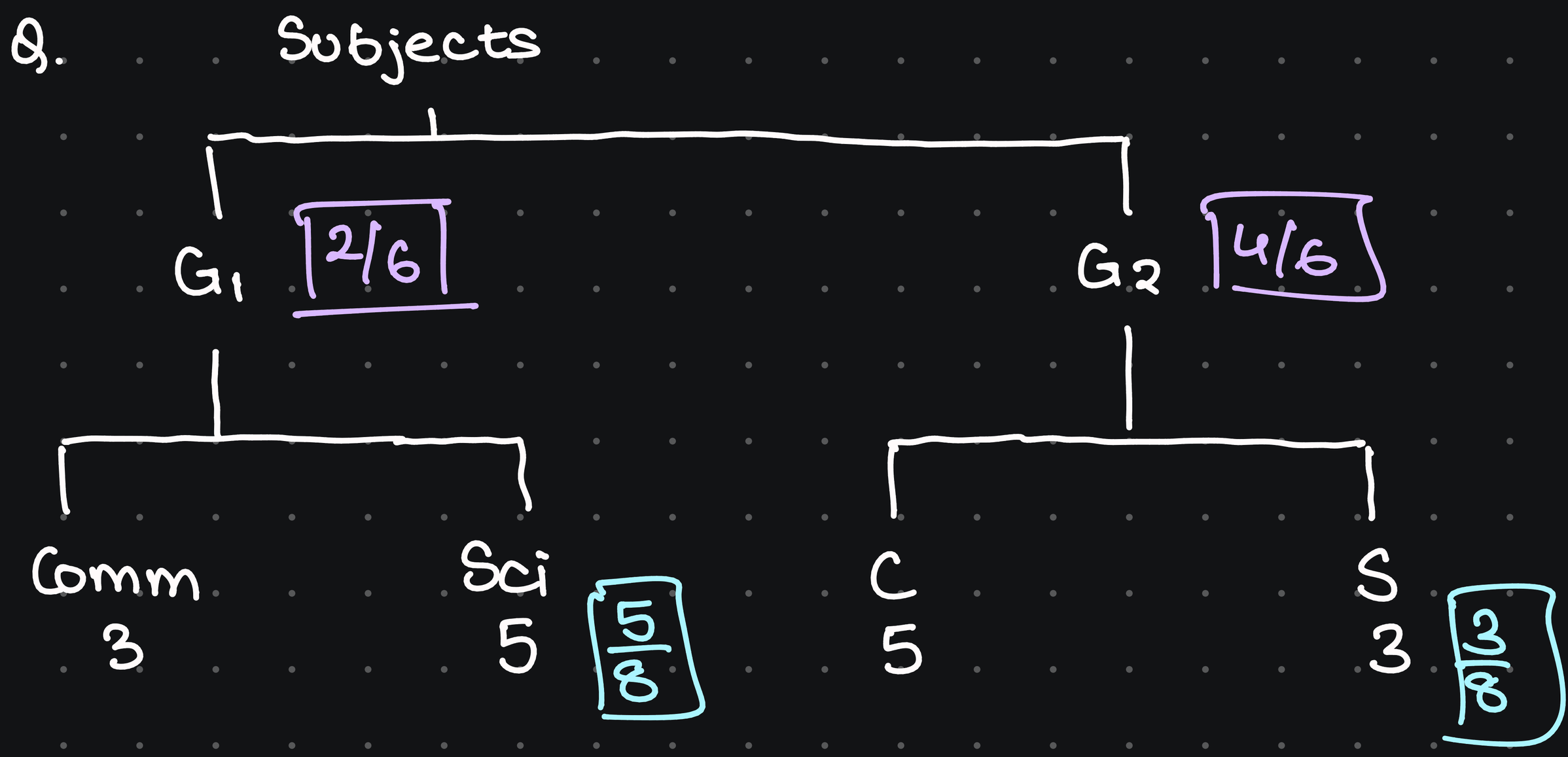


$P(S) = P(S_1 \cap P_1)$ or $P(S_2 \cap P_2)$

$= P(S_1) \cdot P(P_1) + P(S_2) \cdot P(P_2)$

$= \frac{3}{8} \cdot \frac{1}{2} + \frac{6}{8} \cdot \frac{1}{2} = \frac{9}{16}$

Q.



throw a die. If outcome more than 4, G_1
Otherwise G_2.

Then subject selected.

i) P(science) ?

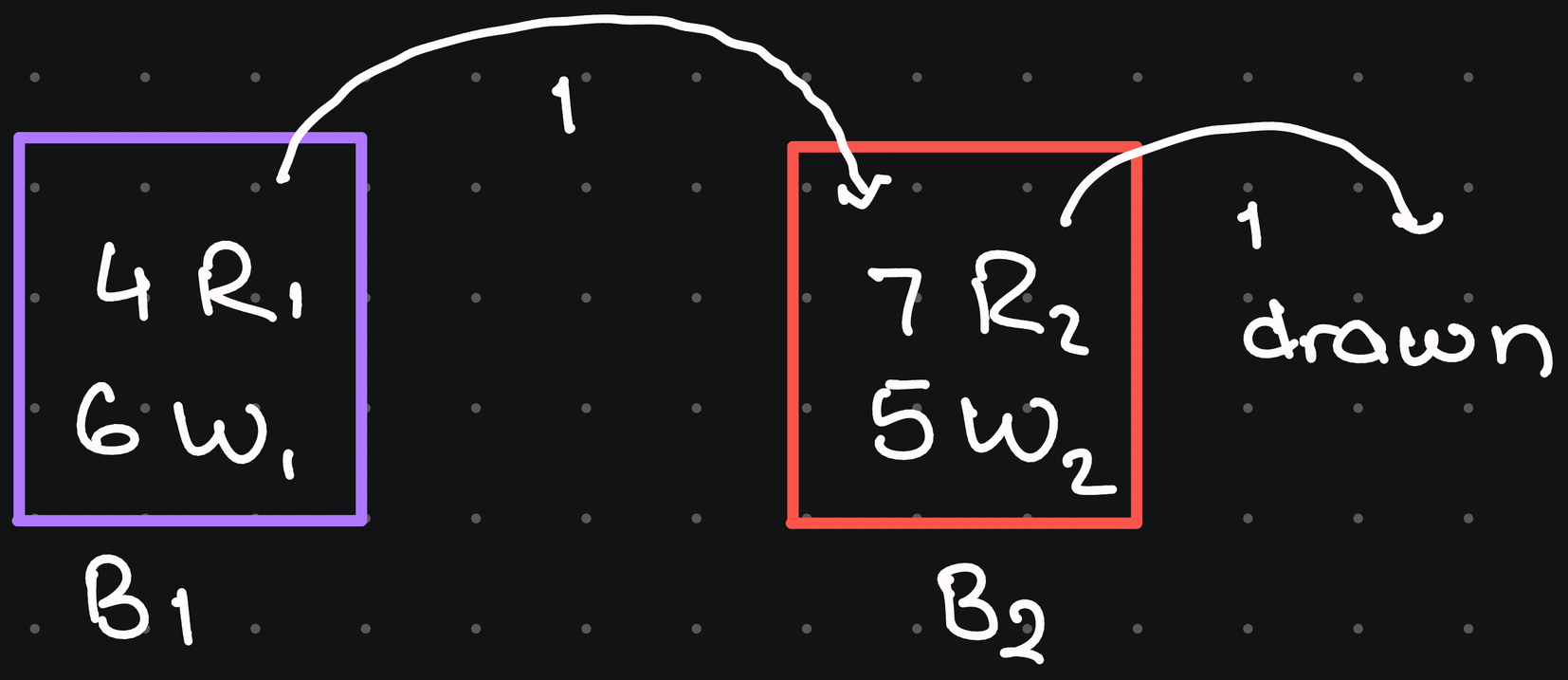
$$\Rightarrow P(\text{Sci.}) = \underset{G_1}{\frac{2}{6} \cdot \frac{5}{8}} \quad \underset{\text{or}}{+} \quad \underset{G_2}{\frac{4}{6} \cdot \frac{3}{8}}$$

$$= \frac{10}{48} + \frac{12}{48}$$

$$= \frac{22}{48}$$

$$= \boxed{\frac{11}{24}}$$

8.



P(ball from B_2 is red)

$\Rightarrow$ P(Red $\cap$ red transf.) or P(Red $\cap$ white t.)

$= P(R_2 \cap R_1) + P(R_2 \cap W_1)$

$= \frac{8}{13} \cdot \frac{4}{10} + \frac{7}{13} \cdot \frac{6}{10}$

$= \frac{74}{130}$

12+1

$= \boxed{\frac{37}{65}}$

Q. A & B 2 events

$P(A) = p$

$P(B) = 1-p$

$P(A \cap B) = (1-p)^2$

$P(A \cap B) = \phi$

disjoint event

For what values of p. A & B are independent?

⇒ for independent events.

$P(A \cap B) = P(A) \cdot P(B)$

$\therefore (1-p)^2 = p(1-p)$

$1-p = p \quad \text{or} \quad 1-p = 0$

$1 = 2p \qquad \therefore p = 1$

$\therefore \boxed{p = \frac{1}{2} \quad \text{or} \quad p = 1}$

a. $\Omega = \{w_1 . w_2 . w_3\}$ sample space

$P(w_1) = k$

$P(w_2) = 2k^2$

$P(w_3) = k^2 + k$.

$k = ?$ Examine if $A = \{w_1, w_2\}$ & $B = \{w_2 . w_3\}$ are independent

$\Rightarrow \quad \Sigma p_i = 1$

$\therefore k + 2k^2 + k^2 + k = 1$

$\therefore 3k^2 + 2k - 1 = 0$

$$\alpha, \beta = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

$\therefore 3k^2 + 3k - k - 1 = 0$

$\therefore 3k(k+1) - 1(k+1) = 0$

$\therefore (k+1)(3k-1) = 0$

$\therefore \underbrace{k = -1}$ / $k = \frac{1}{3}$

cant be negative.

$$\therefore k = \frac{1}{3}$$

$$\therefore P(\omega_1) = k = \frac{1}{3}$$

$$P(\omega_2) = 2k^2 = \frac{2}{9}$$

$$P(\omega_3) = k^2 + k = \frac{1}{9} + \frac{1}{3} = \frac{4}{9}$$

$$P(A) = \{\omega_1, \omega_2\} = \frac{1}{3} + \frac{2}{9} = \frac{5}{9}$$

$$P(B) = \{\omega_2, \omega_3\} = \frac{2}{9} + \frac{4}{9} = \frac{6}{9}$$

if A & B are independent,

$$P(A \cap B) = P(A) \cdot P(B)$$

$$LHS = P(\omega_2) = \frac{2}{9}$$

$$RHS = P(A) \cdot P(B) = \frac{5}{9} \cdot \frac{6^2}{9_3} = \frac{10}{27}$$

$$LHS \neq RHS$$

$\therefore$ not independent

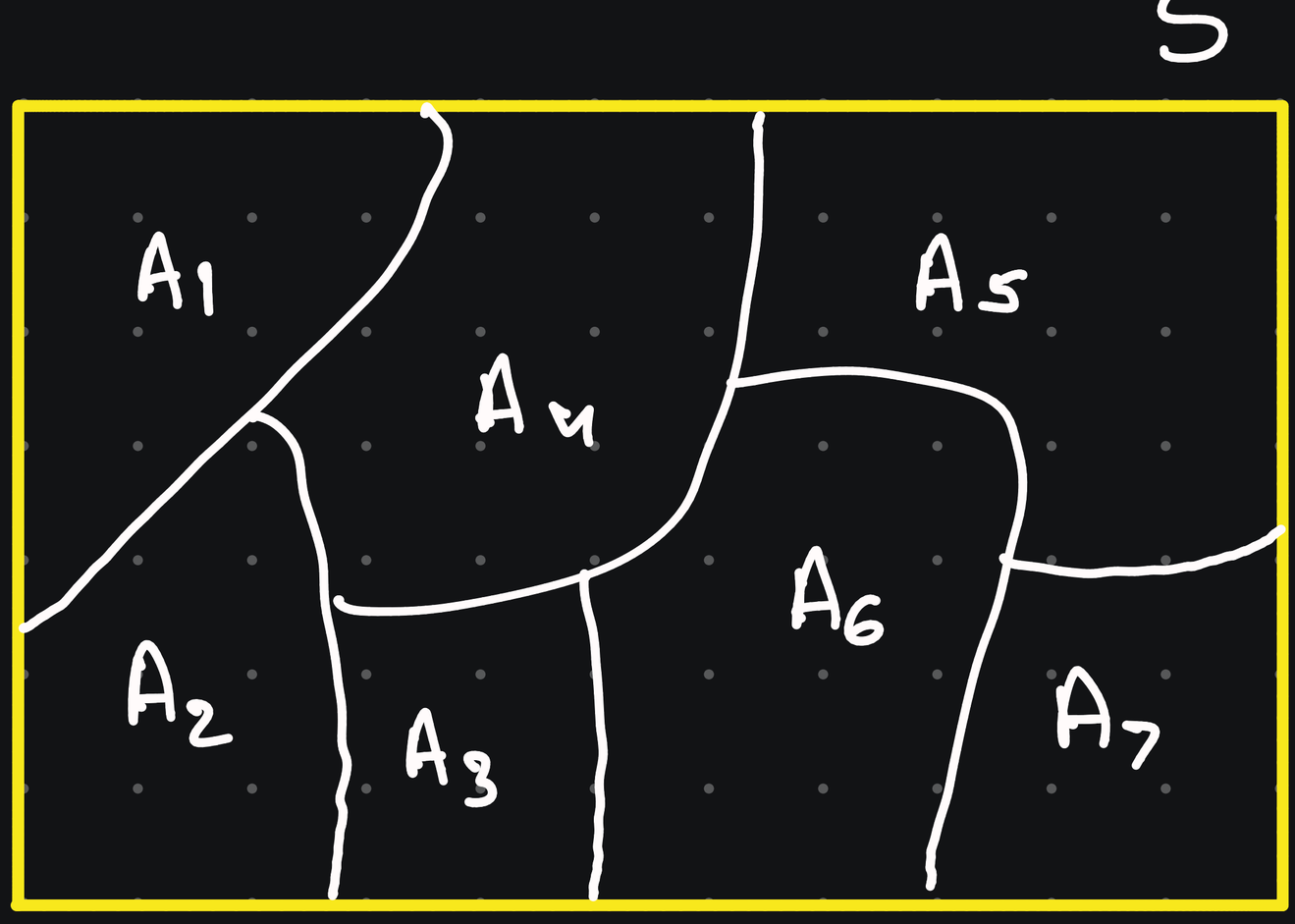
Partition of Sample Space

Let $A_1, A_2, \ldots A_i$ are the events on sample space "S."

They represent **partition on S iff**

i) $A_i \cap A_j = \phi$ for all $i \neq j$ — mutually exclusive

ii) $\bigcup_{i=1}^{n} A_i = S$ — exhaustive



Q. Let A_1, A_2, A_3 denote partitions in ss. "S"

$$P(A_1) = 2(A_2) = 3(A_3)$$

Find $P(A_1 \cup A_2)$

$\Rightarrow$ $P(A_1) = x$

$$\therefore P(A_2) = \frac{P(A_1)}{2} = \frac{x}{2}$$

$$\& \quad P(A_3) = \frac{P(A_1)}{3} = \frac{x}{3}$$

$$P(A_1) + P(A_2) + P(A_3) = 1$$

$$x + \frac{x}{2} + \frac{x}{3} = 1$$

$$6x + 3x + 2x = 6$$

$$11x = 6$$

$$\therefore x = \frac{6}{11}$$

$$P(A_1) = \frac{6}{11}, \quad P(A_2) = \frac{6}{\frac{11}{2}} = \frac{3}{11}, \quad P(A_1 \cap A_2) = 0$$

$$\therefore P(A_1 \cup A_2) = P(A_1) + P(A_2) = \boxed{\frac{9}{11}}$$

Baye's Theorem

Let the events $A_1, A_2, \ldots, A_n$ represent a partition on S.

Let B be any event defined on S.

$P(A_i) \neq 0$ for all i & $P(B) \neq 0$

$$\therefore P(A_i|B) = \frac{P(A_i) \cdot P(B|A_i)}{\sum_{i=1}^{n} \left[P(A_i) \cdot P(B|A_i)\right]}$$

$P(A_i)$ = prior probabilities
(we already know)

$P(A_i|B)$ = posterior probabilities
(calculated later)

Q. Company

Plant I — 70% scooters ($\frac{70}{100}$) — 80% standard quality ($\frac{80}{100}$)

Plant II — 30% scooters — 90% standard quality.

Scooter picked at random, found to be standard quality.

$P(I) = \frac{7}{10}$ $\quad P(II) = \frac{3}{10}$

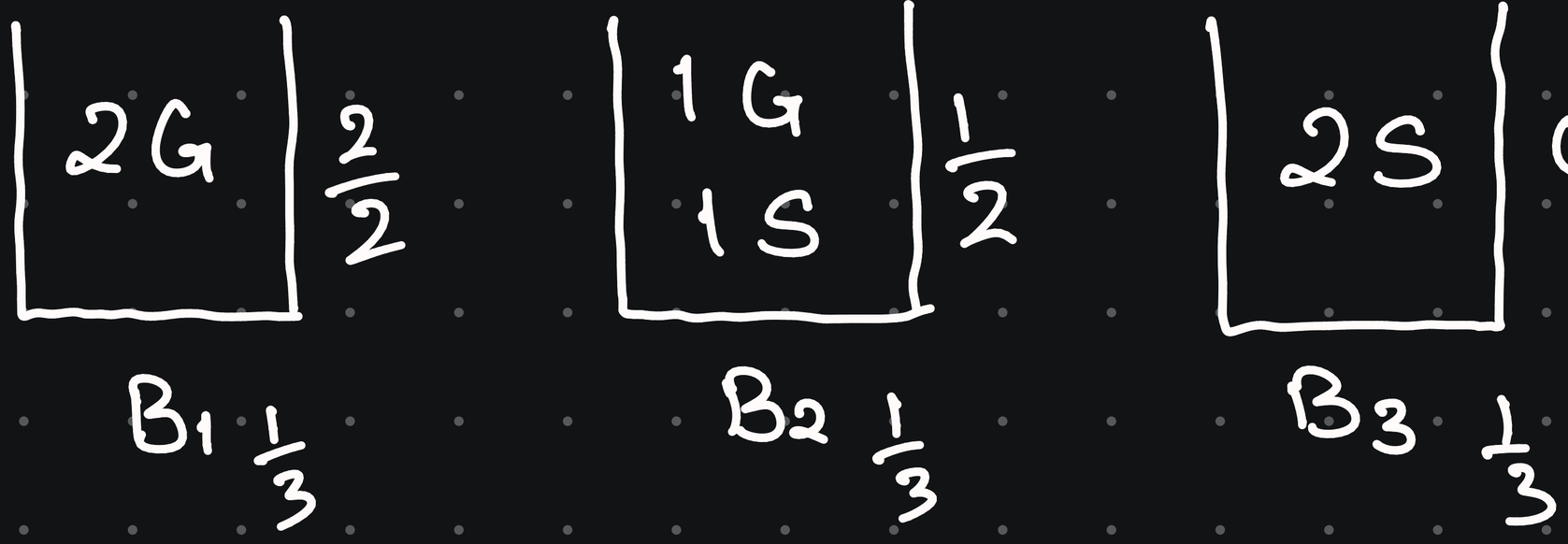
$P(B|I) = \frac{8}{10}$ $\quad P(B|II) = \frac{9}{10}$

$$P(I|B) = \frac{\frac{7}{10} \cdot \frac{8}{10}}{\frac{7}{10} \cdot \frac{8}{10} + \frac{3}{10} \cdot \frac{9}{10}} = \frac{0.56}{0.83}$$

56 $\quad 27$

$$= \boxed{0.6746}$$

8.



Box selected & coin drawn
Coin is gold.
P(other coin in box also gold)?

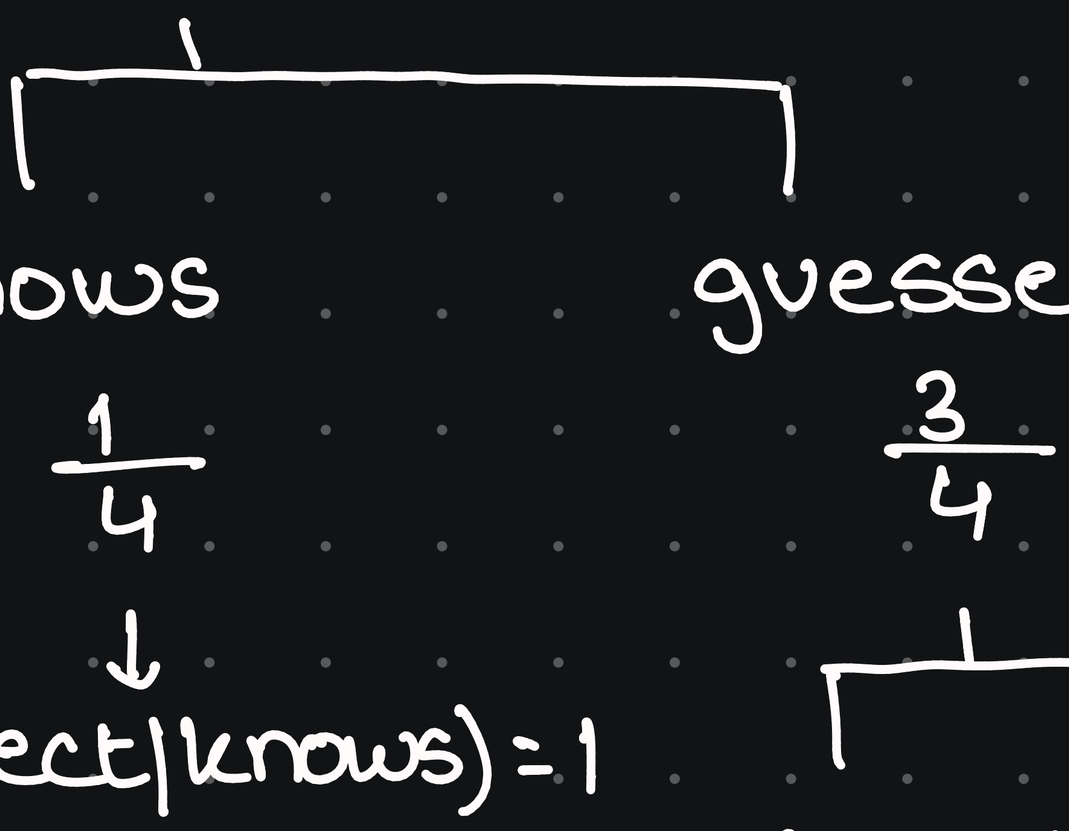
$$\Rightarrow P(B_1|G) = \frac{P(B_1)\cdot P(G|B_1)}{P(B_1)\cdot P(G|B_1) + P(B_2)\cdot P(G|B_2) + P(B_3)\cdot P(G|B_3)}$$

$$= \frac{\frac{1}{3}\cdot\frac{2}{2}}{\frac{1}{3}\cdot\frac{2}{2} + \frac{1}{3}\cdot\frac{1}{2} + 0}$$

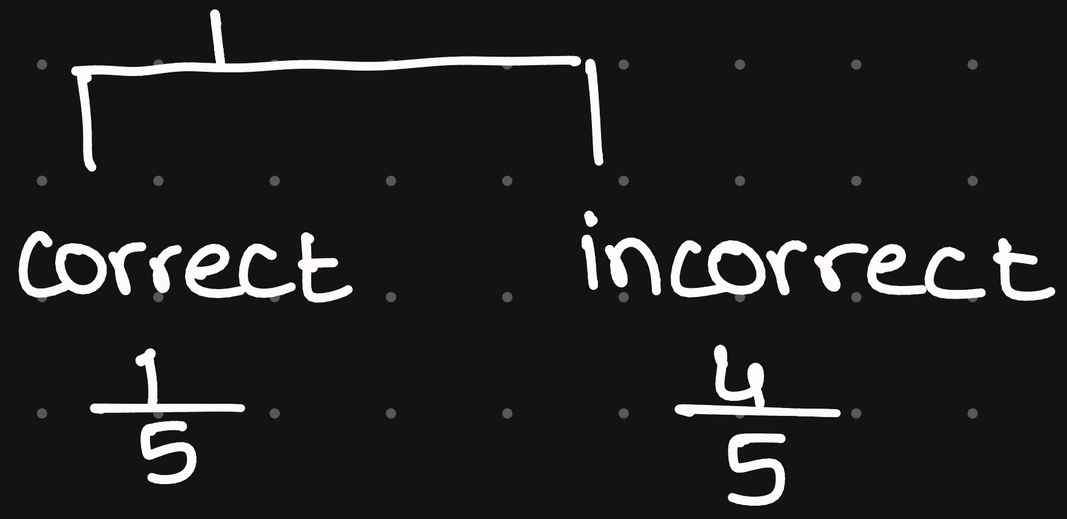
$$= \frac{\frac{2}{6}}{\frac{2}{6} + \frac{1}{6}}$$

$$= \boxed{\frac{2}{3}}$$

9. MCQ test



$P(\text{correct}|\text{knows}) = 1$



P(a student knew the answer given that it is correct)

$$\Rightarrow P(k|c) = \frac{P(k)\cdot P(c|k)}{P(k)\cdot P(c|k) + P(g)\cdot P(c|g)}$$

$$= \frac{\frac{1}{4}\cdot 1}{\frac{1}{4}\cdot 1 + \frac{3}{4}\cdot\frac{1}{5}}$$

$$= \frac{\frac{1}{4}}{\frac{5}{20}}$$

$$= \frac{\frac{1}{4}}{\frac{5}{20}+\frac{3}{20}}$$

$$= \boxed{\frac{5}{8}}$$

True positive → sick diagnosed as sick

e False positive → healthy diagnosed as sick

True negative → healthy diagnosed as healthy

e False negative → sick diagnosed as healthy

$$\text{Sensitivity} = \frac{TP}{TP + FN}$$

ability of test to correctly identify the disease

$$\text{Specificity} = \frac{TN}{TN + FP}$$

Q.

Condition →	Positive	Negative
Test outcome positive	20 TP	180 FP
Test outcome negative	10 FN	1820 TN

compute sensitivity & specificity.

$$\Rightarrow \text{Sensitivity} = \frac{TP}{TP+FN}$$

$$= \frac{20}{20+10}$$

$$= \boxed{\frac{2}{3}} \quad = \boxed{66.67\%}$$

$$\text{Specificity} = \frac{TN}{TN+FP}$$

$$= \frac{1820}{1820+180}$$

$$= \cancel{\frac{1820}{2000}}\ \frac{91}{100}$$

$$= \boxed{0.91} \quad = \boxed{91\%}$$

Concept of Random Variable

Some stats. expts. result in terms of qualitative outcomes & some expts. result in terms of quantitative outcomes.

In quantitative outcomes, mathematical operations can be done directly.
But in qualitative outcomes. mathematical operations cannot be performed directly.
So, we have to assign real numbers to qualitative outcomes. These real nos. are called as random variables.

x	x_1	x_2	x_3	....	x_n
$P(x)$	$P(x_1)$	$P(x_2)$	$P(x_3)$	....	$P(x_n)$

Probability mass function (p.m.f.)

$P(x) \geq 0$

$\Sigma P(x_i) = 1$

discrete probability distribution !

$$E(x) = \Sigma\, x \cdot p(x)$$

expectation

Q. Tossing 3 coins

$S = \{HHH, HHT, HTT, HTH, TTH, THH, THT, TTT\}$

$x \rightarrow$ getting a head

PMF:

x	0	1	2	3
$P(x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

Q.

Car	days	P(x)	
A	23	$\frac{23}{270}$	0.085
B	37	$\frac{37}{270}$	0.137
C	81	$\frac{81}{270}$	0.3
D	53	$\frac{53}{270}$	0.196
E	76	$\frac{76}{270}$	0.281

mean / expectation

$$E(x) = \Sigma x \cdot p(x) = \Sigma p_i x_i$$

Variance

$$V(x) = E(x^2) - [E(x)]^2$$

$$= \Sigma x^2 \cdot p(x) - [\Sigma x \cdot p(x)]^2$$

$$= \Sigma p_i (x_i)^2 - (\Sigma p_i x_i)^2$$

Q.	x_i	$p(x_i)$	$p_i x_i$	x_i^2	$p_i x_i^2$
	1	0·2	0·2	1	0·2
	2	0·3	0·6	4	1·2
	3	0·35	1·05	9	3·15
	4	0·05	0·2	16	0·8
	5	0·1	0·5	25	2.5
		$\Sigma p(x_i) = 1$	2·55		7.85

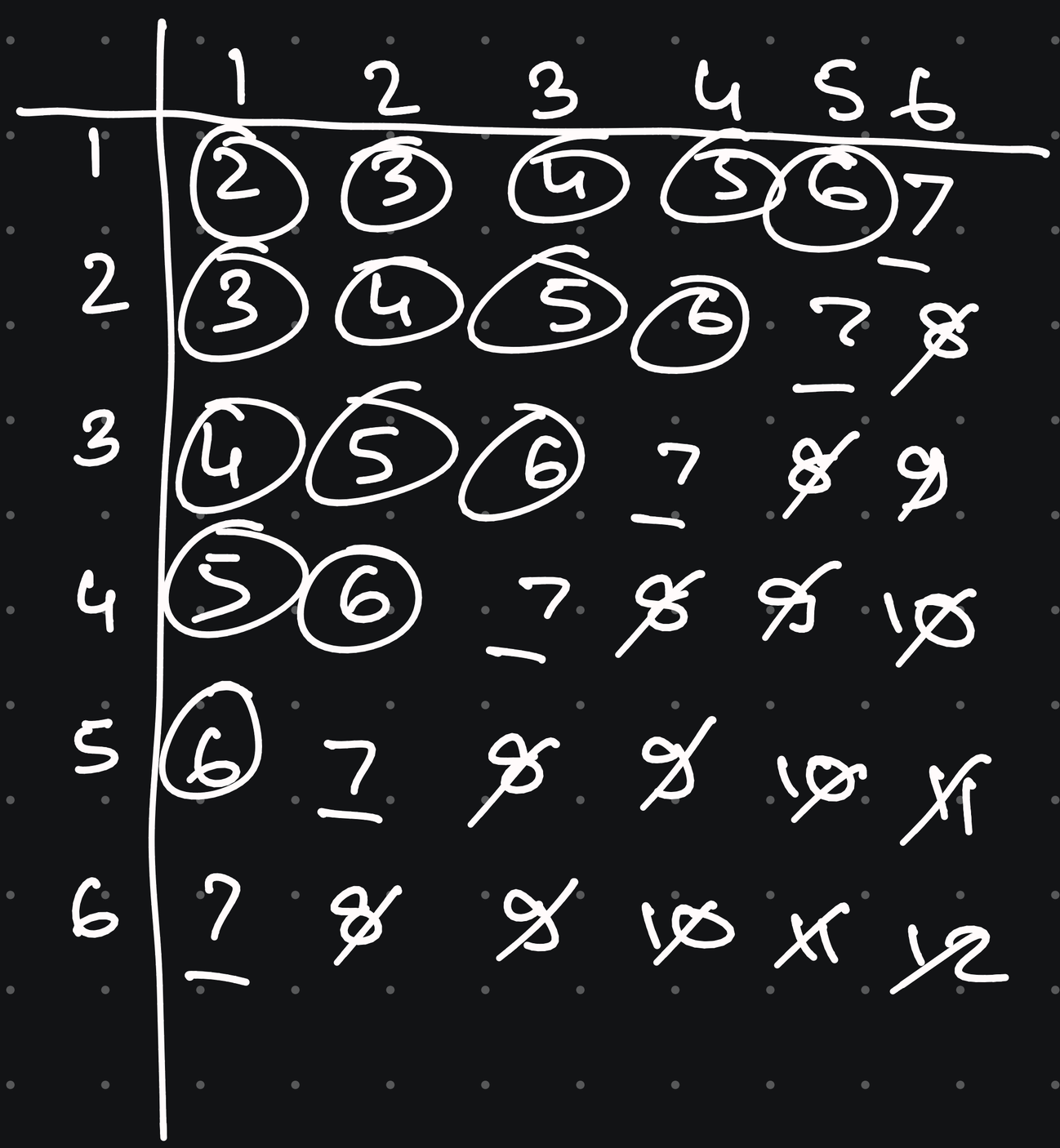
$E(x) = \boxed{2.55}$

$V(x) = 7.85 - 6.5025 = \boxed{1.3475}$

Q. find $E(x)$ & $V(x)$ of 7up-7down
Interpret
How applicable in business?

} fair game (equiprobable)

	$P(x)$	
2-6	$\frac{15}{36}$	$2x$
7	$\frac{6}{36}$	$3x$
7-12	$\frac{15}{36}$	$2x$



x	$P(x)$	$P(x \leq x)$ — cumulative mass function (cmf)
1	0.2	0.2
2	0.3	0.5
3	0.35	0.85 → P(first three are solved)
4	0.05	0.9
5	0.1	1

Q. Cab driver earns net profit of ₹2800 if it rains. Otherwise net profit ₹1500. Windy weather net profit ₹2000. $E(x) = ?$

x		$P(x)$	$p_i x_i$
1500	(not)	0.35	525
2000	(wind)	0.20	400
2800	(rain)	0.45	1260
			2185

$$E(x) = \Sigma p_i x_i = \underline{2185}$$

Q. | 4R
6W |

draw 2 balls

Expectation of red balls?

⇒

x_i	$P(x_i)$	$p_i x_i$
0	$\frac{6C_2}{10C_2} = \frac{1}{3}$	0
1	$\frac{4C_1 \cdot 6C_1}{10C_2} = \frac{8}{15}$	$\frac{8}{15}$
2	$\frac{4C_2}{10C_2} = \frac{2}{15}$	$\frac{4}{15}$
		$\frac{12}{15}$

$$E(\text{red}) = \frac{4}{5} = 0.8$$

Discrete

x_i

p_i

$\Sigma p_i = 1$

P.m.f.

C.m.f.

cmf = $P(x \leq x)$

Continuous

$\int_{-\infty}^{\infty} f(x)\,dx = 1$

Area under functn is 1.

p.d.f.

└ density

c.d.f.

$F(x) = P(x \leq x)$

Q. $f(x) = \frac{3}{4}x(2-x)$.

$0 \leq x \leq 2$

is $f(x)$ p.d.f?

$\Rightarrow f(x)$ is pdf if $\int_0^2 f(x)\,dx = 1$

$$\frac{3}{4}\int_0^2 x(2-x)\,dx = \frac{3}{4}\int_0^2 (2x - x^2)\,dx$$

$$= \frac{3}{4}\left[x^2 - \frac{x^3}{3}\right]_0^2$$

$$= \frac{3}{4}\left[2^2 - \frac{2^3}{3} - 0\right]$$

$$= \frac{3}{4}\left[4^1 - \frac{8^2}{3}\right]$$

$$= 3 - 2$$

$$= 1$$

$\therefore$ $f(x)$ is pdf.

Q. $f(x) = kx^4 \quad -1 \le x \le 0$

$= 0$ otherwise

$f(x)$ is pdf. $k = ?$ $P(x > \frac{-1}{2}) = ?$

$\Rightarrow$ $f(x)$ is pdf if $\int_{-1}^{0} f(x)dx = 1$

$$1 = \int_{-1}^{0} kx^4 dx$$

$$1 = k\left[\frac{x^5}{5}\right]_{-1}^{0}$$

$$1 = k\left[0 - \left(\frac{-1}{5}\right)\right]$$

$$1 = k\left(\frac{1}{5}\right)$$

$$\boxed{k = 5}$$

$$P\left(x > \frac{-1}{2}\right) = \int_{\frac{-1}{2}}^{0} 5x^4 dx$$

$$= 5\int_{\frac{-1}{2}}^{0} x^4 dx = 5\left[\frac{x^5}{5}\right]_{\frac{-1}{2}}^{0} = 0^5 - \left(\frac{-1}{2}\right)^5$$

$$= \boxed{\frac{1}{32}}$$

$$E(ax+b) = a \cdot E(x) + b$$

$$E(x) = \int_0^1 x \cdot f(x)\,dx$$

$$E(x^2) = \int_0^1 x^2 \cdot f(x)\,dx$$

$$V(x) = E(x^2) - [E(x)]^2$$

Q. $f(x) = 3x^2 \quad 0 \le x \le 1$

$= 0$ otherwise

$E(x) = ?$ $V(x) = ?$ $E(3x+2) = ?$

$$\Rightarrow E(x) = \int_0^1 x \cdot 3x^2\,dx = \int_0^1 3x^3\,dx = 3\int_0^1 \frac{x^4}{4}$$

$$= 3\left(\frac{1}{4}\right) = \boxed{\frac{3}{4}}$$

$$E(x^2) = \int_0^1 x^2 \cdot 3x^2\,dx = \int_0^1 3x^4\,dx = 3\int_0^1 \frac{x^5}{5}$$

$$= 3\left(\frac{1}{5}\right) = \boxed{\frac{3}{5}}$$

$$E(3x+2) = 3 \cdot E(x) + 2$$

$$= 3 \cdot \frac{3}{4} + 2$$

$$= \frac{9}{4} + \not{2}\ \frac{8}{4}$$

$$= \boxed{\frac{17}{4}}$$

$$\vartheta(x) = E(x^2) - [E(x)]^2$$

$$= \frac{3}{5} - \frac{9}{16}$$

$$= \boxed{0.0375}$$

$$\boxed{F(x) = \int_{-\infty}^{x} f(t)\,dt}$$

$$f(x) = \int_{0}^{x} 3t^2\,dt \quad = 3\int_{0}^{x} \frac{t^3}{3} \quad = \underline{\underline{x^3}}$$

Q. $F(x) = 0$ if $x < 0$

$= x^2$ if $0 \leq x < \frac{1}{2}$

$= 1 - \frac{3}{2}(1-x)$ if $\frac{1}{2} \leq x \leq 1$

$= 1$ if $x > 1$

Obtain pdf $f(x)$ & find $E(x)$

$\Rightarrow$ **derivative of $F(x) = f(x)$**

ii) $f(x) = \frac{x^2}{dx} = 2x$

iii) $f(x) = \frac{1 - \frac{3}{2}(1-x)}{dx} = \frac{1 - \frac{3}{2} + \frac{3}{2}x}{dx} = \frac{3}{2}$

$\therefore f(x) = 2x, \quad 0 \leq x < \frac{1}{2}$

$= \frac{3}{2}, \quad \frac{1}{2} \leq x \leq 1$

$= 0, \quad x > 1$

$$E(x) = \int_0^1 x f_x \, dx$$

$$E(x) = \int_0^{1/2} x \cdot 2x \, dx + \int_{1/2}^{1} x \frac{3}{2} \, dx$$

$$= 2 \int_0^{0.5} x^2 \, dx + \frac{3}{2} \int_{0.5}^{1} x \, dx$$

$$= 2 \left[\frac{x^3}{3} \right]_0^{1/2} + \frac{3}{2} \left[\frac{x^2}{2} \right]_{1/2}^{1}$$

$$= 2 \left[\frac{0}{3} + \frac{\frac{1}{8}}{3} \right] + \frac{3}{2} \left[\frac{-\frac{1}{4}}{2} + \frac{1}{2} \right]$$

$$= \cancel{2} \left[-0 + \frac{1}{\cancel{2} \times 12} \right] + \frac{3}{2} \left[-\frac{1}{8} + \frac{1}{2} \right]$$

$$= \frac{+1}{12} - \frac{3}{16} + \frac{3}{4}$$

$$= \boxed{\frac{31}{48}}$$

Addith thm / Multiplicath
Bayes thm
Conditional distributh

} proof